

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2004, 09:06:38 ; Search time 2.93651 Seconds  
(without alignments)  
1250.845 Million cell updates/sec

Title: US-09-785-215-19  
Perfect score: 64  
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					Description
Result	Query	Match	Length	DB	ID		
No.	Score						
1	64	100.0	13	2	AAR70250	Aar70250 Pan DR-bi	
2	64	100.0	13	2	AAW22121	Aaw22121 Padre (pa	
3	64	100.0	13	3	AAy52558	Aay52558 Universal	
4	64	100.0	13	3	AAB36289	Aab36289 Promiscuo	
5	64	100.0	13	4	AAB99710	Aab99710 Pan-DR-bi	
6	64	100.0	13	4	AAB73644	Aab73644 Pan-DR bi	
7	64	100.0	13	4	AAB20154	Aab20154 PADRE pep	
8	64	100.0	13	5	AAU80293	Aau80293 Pan DR ep	
9	64	100.0	13	5	ABG31776	Abg31776 Pan DR ep	

10	64	100.0	13	5	AAE26369	Aae26369 PADRE pep
11	64	100.0	13	6	ABP72696	Abp72696 T cell ep
12	64	100.0	13	6	AAO30456	Aao30456 Pan DR ep
13	64	100.0	13	6	AAO30463	Aao30463 Pan DR ep
14	64	100.0	13	7	ABR82484	Abr82484 Universal
15	64	100.0	13	7	ADC81611	Adc81611 Universal
16	64	100.0	55	5	AAU80290	Aau80290 Pan DR ep
17	64	100.0	80	3	AAAY52548	Aay52548 Murine Ig
18	64	100.0	80	7	ADA49413	Ada49413 Multi-epi
19	64	100.0	98	7	ADA49417	Ada49417 Multi-epi
20	64	100.0	106	7	ADA49411	Ada49411 Multi-epi
21	64	100.0	107	7	ADA49409	Ada49409 Multi-epi
22	64	100.0	107	7	ADA49419	Ada49419 Multi-epi
23	64	100.0	118	3	AAAY52541	Aay52541 KappaLAMP
24	64	100.0	119	3	AAAY52542	Aay52542 Murine H2
25	64	100.0	122	3	AAAY52576	Aay52576 Amino aci
26	64	100.0	123	7	ADA49425	Ada49425 Multi-epi
27	64	100.0	130	7	ADA49415	Ada49415 Multi-epi
28	64	100.0	132	3	AAAY52575	Aay52575 Amino aci
29	64	100.0	136	3	AAAY52543	Aay52543 Murine H2
30	64	100.0	138	3	AAAY52539	Aay52539 Murine Ii
31	64	100.0	144	7	ADA49445	Ada49445 Multi-epi
32	64	100.0	147	7	ADA49447	Ada49447 Multi-epi
33	64	100.0	148	7	ADA49443	Ada49443 Multi-epi
34	64	100.0	152	3	AAAY52547	Aay52547 Murine Ig
35	64	100.0	157	6	AAO30484	Aao30484 Human TNF
36	64	100.0	157	6	AAO30485	Aao30485 Human TNF
37	64	100.0	157	6	AAO30483	Aao30483 Human TNF
38	64	100.0	157	6	AAO30481	Aao30481 Human TNF
39	64	100.0	157	7	ADA49433	Ada49433 Multi-epi
40	64	100.0	160	6	AAO30482	Aao30482 Human TNF
41	64	100.0	162	6	AAO30473	Aao30473 Human TNF
42	64	100.0	164	3	AAAY52546	Aay52546 Murine Ig
43	64	100.0	165	6	AAO30475	Aao30475 Human TNF
44	64	100.0	165	6	AAO30469	Aao30469 Human TNF
45	64	100.0	167	6	AAO30468	Aao30468 Human TNF

# ALIGNMENTS

## RESULT 1

AAR70250

ID AAR70250 standard; peptide; 13 AA.

XX

AC AAR70250;

XX

DT 25-MAR-2003 (revised)

DT 13-NOV-1995 (first entry)

XX

DE Pan DR-binding peptide activator of T cells.

XX

KW T cell; MHC; class II; major histocompatibility complex II; immunogen;  
KW activation.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /note= "D-form residue"  
 FT Misc-difference 13  
 FT /note= "D-form residue"  
 XX  
 PN WO9507707-A1.  
 XX  
 PD 23-MAR-1995.  
 XX  
 PF 14-SEP-1994; 94WO-US010368.  
 XX  
 PR 14-SEP-1993; 93US-00121101.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Sette A, Gaeta FC, Grey HM, Sidney J, Alexander JL;  
 XX  
 DR WPI; 1995-131178/17.  
 XX  
 PT Inhibiting or inducing an immune response using Pan DR-binding peptide(s)  
 PT - to prepare a compsn. used to treat allo:graft rejection, allergic  
 PT response and auto:immunity and as a vaccine component.  
 XX  
 PS Claim 3; Page 51; 59pp; English.  
 XX  
 CC Peptides capable of binding MHC class II molecules via antigen binding  
 CC sites (pan DR-binding peptides) were synthesised. Peptide AAR70250 is  
 CC capable of binding MHC class II molecules causing T cell activation. This  
 CC can be used as a vaccine component. The peptide can also be used in  
 CC combination with CTL peptides to enhance a CTL response. (Updated on 25-  
 CC MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

# RESULT 2

AAW22121

ID AAW22121 standard; peptide; 13 AA.

XX

AC AAW22121;

XX

DT 13-MAR-1998 (first entry)

XX

DE Padre (pan-DR binding) peptide 6.

XX

KW PADRE; pan-DR binding; immune response; antigenic determinant; treatment;  
 KW tumour; infection.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1

FT /note= "D-form residue"

FT Misc-difference 13

FT /note= "D-form residue"

XX

PN WO9726784-A1.

XX

PD 31-JUL-1997.

XX

PF 23-JAN-1997; 97WO-US001041.

XX

PR 24-JAN-1996; 96US-0010510P.

XX

PA (CYTE-) CYTEL CORP.

XX

PI Alexander JL, Defrees S, Sette A;

XX

DR WPI; 1997-393272/36.

XX

PT Composition for eliciting immune response to non-protein determinant -

PT comprises the determinant and a pan-DR binding peptide, used for

PT prevention and treatment of tumours and infections.

XX

PS Claim 23; Page 74; 87pp; English.

XX

CC This PADRE (pan-DR binding) peptide is used in a composition for

CC eliciting an immune response to a non-protein antigenic determinant. The

CC composition comprises of the PADRE peptide covalently linked to the non-

CC protein antigenic determinant. The composition is used to induce a

CC therapeutic or prophylactic response, particularly to selected

CC polysaccharide antigens associated with tumours or infectious agents. It

CC provides a high level, long-lasting IgG immune response. The composition

CC is also used to produce monoclonal antibodies which are potentially

CC useful as therapeutic and diagnostic agents. The composition can be used

CC to diagnose susceptibility of a patient to treatment with the non-protein

CC antigenic determinant or to predict subjects at risk from developing

CC chronic infections. PADRE peptides have broader specificity and higher

CC affinity than known DR-binding peptides. They are powerful inhibitors of

CC the proliferative response of human T cells restricted by at least 6

CC different DR molecules, and act as helper epitopes of in vivo induction

CC of cytotoxic T cells and antibody production

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

RESULT 3

AA52558

ID AAY52558 standard; peptide; 13 AA.

XX

AC AAY52558;

XX

DT 28-FEB-2000 (first entry)

XX

DE Universal helper T epitope, pan DR epitope (PADRE).

XX

KW Chimeric; Ii protein; pan DR epitope; expression vector; promoter;  
KW major histocompatibility complex; MHC; targetting; peptide; epitope;  
KW antigen; presentation; class I; cytosolic pathway; endoplasmic reticulum;  
KW class II; extracellular antigen; endocytic pathway; helper T lymphocyte;  
KW HTL; universal epitope; cytotoxic T lymphocyte; CTL; immune response;  
KW immunogenicity; assay; vaccine; immunity; infection; pathogen; virus;  
KW HIV; HBV; HCV; hepatitis B; hepatitis C; bacterium; protozoan;  
KW tumour cell; autoimmune disease; activation; antiviral; antimalarial;  
KW immunoprotective.

XX

OS Synthetic.

XX

PN WO9958658-A2.

XX

PD 18-NOV-1999.

XX

PF 13-MAY-1999; 99WO-US010646.

XX

PR 13-MAY-1998; 98US-00078904.

PR 15-MAY-1998; 98US-0085751P.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
PI Chesnut RW;

XX

DR WPI; 2000-039103/03.

DR N-PSDB; AAZ38679.

XX

PT Expression vectors encoding major histocompatibility targeting sequence,  
PT used as, e.g. tumor vaccines.

XX

PS Claim 9; Page 80; 130pp; English.

XX

CC This sequence represents a universal helper T epitope, pan DR epitope  
CC (PADRE), DNA encoding which is used to construct fusion genes used in  
CC exemplifications of the present invention. The invention relates to a  
CC novel expression vector comprising a promoter operably linked to a fusion  
CC gene encoding a major histocompatibility complex (MHC) targetting  
CC sequence, and two or more heterologous peptide epitopes. The MHC  
CC targetting sequence may be a class I targetting sequence, which directs  
CC an MHC class I epitope to a cytosolic pathway or to the endoplasmic  
CC reticulum, or an MHC class II targetting sequence, which directs  
CC extracellular antigens to enter the endocytic pathway to be processed  
CC into antigen peptides for presentation on MHC class II molecules. The  
CC heterologous epitopes may comprise either helper T lymphocyte (HTL)  
CC epitopes, or a cytotoxic T lymphocyte (CTL) epitope and a universal HTL  
CC epitope such as a pan DR epitope (PADRE). The vectors are useful for

CC stimulating an immune response in vivo, as well as for use in assaying  
CC the human immunogenicity of a human T cell peptide epitope in vivo in a  
CC non-human mammal. They provide a nucleic acid vaccine for enhancing  
CC immunity against infectious pathogens, such as viruses (e.g., HIV,  
CC hepatitis B (HBV and hepatitis C (HCV)) bacteria, protozoa (e.g.,  
CC Plasmodium falciparum, the cause of malaria) and also tumour cells and  
CC autoimmune diseases. Universal MHC class II epitopes are advantageously  
CC combined with other MHC class I and class II epitopes to increase the  
CC number of cells that are activated in response to a given antigen and  
CC provide a broader population coverage of MHC-reactive alleles  
XX  
SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

#### RESULT 4

AAB36289

ID AAB36289 standard; peptide; 13 AA.

XX

AC AAB36289;

XX

DT 26-FEB-2001 (first entry)

XX

DE Promiscuous T helper epitope SEQ ID NO: 65.

XX

KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;  
KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX

OS Synthetic.

XX

PN WO200065058-A1.

XX

PD 02-NOV-2000.

XX

PF 19-APR-2000; 2000WO-DK000205.

XX

PR 23-APR-1999; 99DK-00000552.

PR 06-MAY-1999; 99US-0132811P.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Klysner S;

XX

DR WPI; 2000-672791/65.

XX

PT Down-regulating interleukin 5 (IL-5) activity in humans by administering  
PT IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or  
PT amelioration of asthma or other chronic allergic conditions.

XX

PS Disclosure; Page 168; 172pp; English.

XX  
CC The present invention is concerned with methods of treating asthma,  
CC eosinophilia, allergic rhinitis and other allergic diseases. These  
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
CC proteins and their coding sequences to down-regulate IL-5 activity and  
CC thus reduce eosinophil numbers. The allergic diseases may be treated  
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
CC it is possible that they may be used in the treatment of cancer and  
CC helminthic infections  
XX  
SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 5

AAB99710

ID AAB99710 standard; peptide; 13 AA.

XX

AC AAB99710;

XX

DT 06-SEP-2001 (first entry)

XX

DE Pan-DR-binding peptide (PADRE) SEQ ID NO:31.

XX

KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;  
KW cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;  
KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;  
KW immunotherapy; immune response.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200141741-A1.

XX

PD 14-JUN-2001.

XX

PF 13-DEC-2000; 2000WO-US034318.

XX

PR 13-DEC-1999; 99US-0170448P.

PR 05-APR-2000; 2000US-00543608.

PR 30-MAY-2000; 2000US-00583200.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;  
PI Chesnut R;

XX

DR WPI; 2001-381489/40.

XX

PT Compositions for use in a vaccine for treating, e.g., breast, lung and

PT colon cancer comprises at least one peptide that comprises an isolated  
PT epitope of a tumor-associated antigen.

XX

PS Example 7; Page 48; 86pp; English.

XX

CC The present invention describes a composition (I) comprising at least one  
CC peptide that comprises an isolated, prepared epitope consisting of a  
CC sequence selected from 25 short amino acid sequences given in AAB99680 to  
CC AAB99704. Also described are: (1) a composition (II) comprising one or  
CC more peptides, and further comprising at least two epitopes selected from  
CC the 25 short amino acid sequences (as above), where each of the one or  
CC more peptides comprise less than 50 contiguous amino acids that have 100%  
CC identity with a native peptide sequence; and (2) a vaccine composition  
CC (III) comprising an epitope selected from the 25 short amino acid  
CC sequences (as above) and a pharmaceutical excipient. (I) has cytostatic  
CC and immunomodulatory activities and can be used in vaccine production and  
CC immunotherapy. The peptide epitope compositions (I)-(II) are useful for  
CC monitoring an immune response to a tumour associated antigen or when one  
CC or more peptides are combined to create a vaccine (III) that stimulates  
CC the cellular arm of the immune system. In particular, the vaccine  
CC mediates immune responses against tumours in individuals who bear an  
CC allele of the human leukocyte antigen (HLA)-A2 supertype and improve the  
CC standard of care for patients being treated for breast, colon, or lung  
CC cancer. The present sequence represents a pan-DR-binding peptide (PADRE)  
CC sequence, which is used in an example from the present invention

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 4; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

#### RESULT 6

AAB73644

ID AAB73644 standard; peptide; 13 AA.

XX

AC AAB73644;

XX

DT 11-SEP-2001 (first entry)

XX

DE Pan-DR binding peptide, PADRE.

XX

KW Cellular vaccine; antigen-presenting cell; APC;

KW cell-surface molecule density; major histocompatibility complex; MHC;

KW antigen-specific T-cell expansion; tumour; cancer; viral infection;

KW parasitic infection; Pan-DR binding peptide; PADRE.

XX

OS Unidentified.

XX

PN WO200136978-A1.

XX

PD 25-MAY-2001.



XX  
 PF 15-NOV-2000; 2000WO-US042213.  
 XX  
 PR 15-NOV-1999; 99US-0165428P.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schultze JL, Vonderheide RH, Nadler LM, Maecker B;  
 PI Von Bergwelt- Baildon M;  
 XX  
 DR WPI; 2001-343909/36.  
 XX  
 PT Determining surface density of molecules on antigen-presenting cells,  
 PT useful for quality control of cellular vaccines for treatment of tumors  
 PT and infections comprises measuring density ratios.  
 XX  
 PS Disclosure; Page 21; 47pp; English.  
 XX  
 CC The invention relates to a method of determining the surface density of a  
 CC cell-surface molecule on a primary or artificial antigen-presenting cell  
 CC (APC). The method comprises determining the cell surface area of the  
 CC APCs, determining the absolute amount of cell surface molecule on the  
 CC surface of the APCs, and calculating the ratio of the amount of the cell  
 CC surface molecule to the APC surface area as a measure of cell surface  
 CC molecule density. The invention also encompasses a kit for determining  
 CC the cell-surface density of a complex of peptide and MHC (major  
 CC histocompatibility complex) protein/peptide complex on an APC;  
 CC identifying an agent that increases persistence of the MHC/peptide  
 CC complex on an APC surface by culturing APCs in presence of test compound  
 CC and measuring the time of persistence relative to an untreated control;  
 CC and determining if the APC carries a therapeutically adequate amount of  
 CC peptide by determining whether the cell-surface density of the  
 CC MHC/peptide complex is 100 molecules/square micrometre or more. The  
 CC method is used to determine if the amount of MHC/antigenic peptide  
 CC complex present on the surface of an APC is sufficient for therapeutic  
 CC use of the cells as cellular vaccines or for ex vivo expansion of antigen  
 CC -specific T cells for subsequent return to the patient, particularly for  
 CC the treatment of tumours or viral or parasitic infections. The method can  
 CC also be used to identify agents that increase production of MHC/peptide  
 CC complexes on cells, (which increases the therapeutic potential of the  
 CC cells) by treatment before and/or during APC administration. The method  
 CC allows identification of therapeutically useful antigen-pulsed APC  
 CC optimisation of conditions for their production, and control of APC  
 CC quality. Sequences AAB73643-AAB73648 represent peptides used in MHC  
 CC binding studies and for the generation of peptide-specific cytotoxic T-  
 CC lymphocytes  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

RESULT 7

AAB20154

ID AAB20154 standard; peptide; 13 AA.

XX

AC AAB20154;

XX

DT 30-APR-2001 (first entry)

XX

DE PADRE peptide.

XX

KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;

KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;

KW cardiant; PADRE; pan DE epitope.

XX

OS Synthetic.

XX

PN WO200105820-A2.

XX

PD 25-JAN-2001.

XX

PF 20-JUL-2000; 2000WO-DK000413.

XX

PR 20-JUL-1999; 99DK-00001014.

PR 26-JUL-1999; 99US-0145275P.

XX

PA (MEBI-) M & E BIOTECH AS.

XX

PI Halkier T, Mouritsen S, Klysner S;

XX

DR WPI; 2001-112680/12.

XX

PT Increasing the muscle mass of animals used in meat production by down  
PT regulating growth differentiation factor 8 (GDF-8) activity in the animal  
PT through induction of anti-GDF-8 antibody production.

XX

PS Disclosure; Page 15; 110pp; English.

XX

CC The present sequence is that of a PADRE (pan DE epitope) peptide which  
CC acts as a T-cell epitope and is capable of binding a large proportion of  
CC major histocompatibility complex class II molecules. It is an object of  
CC the invention to produce a recombinant therapeutic vaccine that is  
CC capable of effecting down-regulation of growth differentiation factor 8  
CC (GDF-8) in order to increase the muscle growth rate of farm animals. The  
CC vaccines (see AAB20145-53) are capable of breaking autotolerance against  
CC autologous GDF-8. They comprise the C-terminal portion of human GDF-8 in  
CC which a portion of the native sequence is replaced by a T-cell epitope  
CC such as PADRE, with minimal disturbance of the authentic 3-dimensional  
CC structure of the protein. Down-regulation of GDF-8 activity can increase  
CC muscle mass by up to at least 45% in cattle, pigs and poultry used for  
CC meat production, reducing the need for antibiotic feed-additives. Anti-  
CC GDF8 vaccines can be used to treat human diseases such as cancer cachexia  
CC where muscle atrophy is pronounced and for patients suffering from acute  
CC and chronic heart failure

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 8

AAU80293

ID AAU80293 standard; peptide; 13 AA.  
XX  
AC AAU80293;  
XX  
DT 30-JUL-2002 (first entry)  
XX  
DE Pan DR epitope peptides (PADRE) epitope #2.  
XX  
KW IgE; allergy; antiallergic; immunosuppressive; antianaphylactic;  
KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;  
KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;  
KW heavy chain C domain; PADRE; Pan DR epitope peptide.  
XX  
OS Synthetic.  
XX  
PN WO200220038-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 06-SEP-2001; 2001WO-DK000579.  
XX  
PR 06-SEP-2000; 2000DK-00001326.  
PR 15-SEP-2000; 2000US-0232831P.  
XX  
PA (PHAR-) PHARMEXA AS.  
XX  
PI Klysner S, Von Hoegen P, Voldborg B, Gautam A;  
XX  
DR WPI; 2002-383033/41.  
DR N-PSDB; ABK51139.  
XX  
PT Inducing immune response against autologous immunoglobulin E in an  
PT animal, by effecting simultaneous presentation of cytotoxic T lymphocyte  
PT epitope an/or B-cell epitope derived from the immunoglobulin.  
XX  
PS Disclosure; Page 31; 151pp; English.  
XX  
CC This invention relates to a novel method for inducing an immune response  
CC against autologous immunoglobulin E (IgE) in an animal. The method  
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte  
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell  
CC epitope (TH epitope) which is foreign to the animal, by antigen  
CC presenting cells (APCs) of the animal's immune system. The epitopes of  
CC the invention may be used as a vaccine against allergic diseases. The  
CC method of the invention is useful for inducing an immune response against  
CC autologous IgE in an animal, which is useful for downregulating

CC autologous IgE in the animal. This method is useful in the prevention and  
CC treatment of allergic diseases such as anaphylaxis, allergic rhinitis,  
CC asthma and atopic dermatitis. The present sequence represents human Pan  
CC DR epitope (PADRE) #2 used in the method of the invention  
XX  
SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 9

ABG31776

ID ABG31776 standard; peptide; 13 AA.

XX

AC ABG31776;

XX

DT 03-DEC-2002 (first entry)

XX

DE Pan DR epitope (PADRE) peptide.

XX

KW Immunogen; B-cell epitope; cytotoxic T lymphocyte; CTL; TH epitope;

KW T helper cell epitope; virtual lymph node device; pan DR epitope; PADRE.

XX

OS Synthetic.

XX

PN WO200266056-A2.

XX

PD 29-AUG-2002.

XX

PF 19-FEB-2002; 2002WO-DK000112.

XX

PR 19-FEB-2001; 2001WO-DK000113.

PR 20-FEB-2001; 2001US-00785215.

PR 20-AUG-2001; 2001DK-00001231.

PR 22-OCT-2001; 2001US-0337543P.

XX

PA (PHAR-) PHARMEXA AS.

XX

PI Nielsen KG, Koefoed P;

XX

DR WPI; 2002-706932/76.

XX

PT Novel immunogen useful for immunizing an animal, has an activated  
PT polyhydroxypolymer backbone to which is attached an antigenic determinant  
PT including a B cell epitope and another determinant including a T-helper  
PT epitope.

XX

PS Disclosure; Page 17; 52pp; English.

XX

CC The invention relates to an immunogen comprising at least one first

CC antigenic determinant that includes at least one B-cell epitope and/or at

CC least one cytotoxic T lymphocyte (CTL) epitope, and at least one second  
 CC antigenic determinant that includes a T helper cell epitope (TH epitope),  
 CC where each of the first and second antigenic determinants are coupled to  
 CC an activated polyhydroxypolymer carrier. The invention also relates to an  
 CC immunogenic composition for raising an immune response against an antigen  
 CC in a mammal, including a human. The immunogen or immunogenic composition  
 CC contained in a virtual lymph node (VLN) device is useful for immunising  
 CC an animal, including a human, against an antigen of choice, where the  
 CC antigen shares at least one first antigenic determinant with the  
 CC immunogen. This sequence represents a pan DR epitope (PQDRE) peptide  
 CC which is a T helper cell epitope used in synthesis of an immunogen of the  
 CC invention  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 5; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

# RESULT 10

AAE26369

ID AAE26369 standard; peptide; 13 AA.

XX

AC AAE26369;

XX

DT 13-DEC-2002 (first entry)

XX

DE PADRE peptide.

XX

KW Human; immune response; T-helper cell epitope; chitosan; CTL response;

KW vaccine; prostate cancer; breast cancer; cytostatic; immunostimulant.

XX

OS Unidentified.

XX

PN WO200234287-A2.

XX

PD 02-MAY-2002.

XX

PF 26-OCT-2001; 2001WO-DK000705.

XX

PR 27-OCT-2000; 2000DK-00001606.

PR 03-NOV-2000; 2000US-0245166P.

PR 18-JUN-2001; 2001DK-00000936.

XX

PA (PHAR-) PHARMEXA AS.

XX

PI Beier AM, Gautam A, Mouritsen S;

XX

DR WPI; 2002-463339/49.

XX

PT Inducing or enhancing an immune response against an antigen, particularly  
 PT cytotoxic T-lymphocyte responses, for treating or ameliorating prostate

PT or breast cancer, comprises administering the antigen formulated with  
PT chitosan.

XX

PS Disclosure; Page 32; 97pp; English.

XX

CC The invention relates to a method for inducing or enhancing an immune  
CC response against a polypeptide antigen in an animal, including human. The  
CC method comprises administering the polypeptide antigen or at least one  
CC variant which includes at least one first T-helper cell epitope that is  
CC foreign to the animal (foreign TH epitope) and is formulated with  
CC chitosan. The polypeptide antigen is weakly immunogenic or non-  
CC immunogenic. The invention is used as vaccine. The chitosan and  
CC polypeptide antigen or its variant are useful in the preparation of an  
CC immunogenic composition for inducing or enhancing an immune response,  
CC particularly CTL response, against the polypeptide or protein antigen.  
CC The method for inducing or enhancing an immune response is useful in  
CC treating or ameliorating cancer, e.g. prostate or breast cancer. The  
CC present sequence is PADRE peptide used to illustrate the method of the  
CC invention

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 5; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

#### RESULT 11

ABP72696

ID ABP72696 standard; peptide; 13 AA.

XX

AC ABP72696;

XX

DT 11-JUN-2003 (first entry)

XX

DE T cell epitope PADRE.

XX

KW PADRE; epitope; amyloid precursor protein; APP; beta amyloid; vaccine;  
KW genetic immunisation; nootropic; neuroprotective; Alzheimer's disease.

XX

OS Synthetic.

XX

PN WO2003015812-A2.

XX

PD 27-FEB-2003.

XX

PF 20-AUG-2002; 2002WO-DK000547.

XX

PR 20-AUG-2001; 2001DK-00001231.

PR 22-OCT-2001; 2001US-0337543P.

PR 16-APR-2002; 2002DK-00000558.

PR 16-APR-2002; 2002US-0373027P.

XX

PA (PHAR-) PHARMEXA AS.  
 XX  
 PI Rasmussen PB, Jensen MR, Nielsen KG, Koefoed P, Degan FD;  
 XX  
 DR WPI; 2003-312718/30.  
 XX  
 PT Novel analog of amyloid precursor protein or beta amyloid for treating  
 PT Alzheimer's disease, has amyloid precursor protein/beta amyloid  
 PT incorporating B-cell epitope of amyloid protein and foreign T-helper  
 PT epitope.  
 XX  
 PS Disclosure; Page 38; 122pp; English.  
 XX  
 CC The present sequence is that of the pan DR epitope (PADRE), an artificial  
 CC T cell epitope which is capable of binding a large proportion of MHC  
 CC Class II molecules. The invention provides methods for compositions for  
 CC combatting diseases characterised by deposition of amyloid, such as  
 CC Alzheimer's disease. Immunisation is preferably effected by  
 CC administration of analogues of autologous amyloid precursor protein (APP)  
 CC or beta amyloid (Abeta), the analogues being capable of inducing antibody  
 CC production against the autologous amyloidogenic polypeptides. Especially  
 CC preferred as an immunogen is autologous Abeta which has been modified by  
 CC introduction of one or a few foreign, immunodominant and promiscuous T-  
 CC cell epitopes, such as PADRE. Genetic immunisation against APP or Abeta  
 CC and vaccination using live vaccines are also provided  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 6; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

# RESULT 12

AAO30456

ID AAO30456 standard; peptide; 13 AA.

XX

AC AAO30456;

XX

DT 22-SEP-2003 (first entry)

XX

DE Pan DR epitope peptide (PADRE) #1.

XX

KW Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;  
 KW tumour necrosis factor alpha; gene therapy; arthritis; epitope; PADRE;  
 KW pan DR epitope peptide.

XX

OS Unidentified.

XX

PN WO2003042244-A2.

XX

PD 22-MAY-2003.

XX

PF 15-NOV-2002; 2002WO-DK000764.

XX

PR 16-NOV-2001; 2001DK-00001702.

PR 16-NOV-2001; 2001US-0331575P.

XX

PA (PHAR-) PHARMEXA AS.

PA (KLYS/) KLYSNER S.

PA (NIEL/) NIELSEN F S.

PA (BRAT/) BRATT T.

PA (VOLD/) VOLDBORG B.

PA (MOUR/) MOURITSEN S.

XX

PI Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;

XX

DR WPI; 2003-449558/42.

DR N-PSDB; AAL61292.

XX

PT New immunogenic analogue of a polymeric protein, useful for preparing a  
PT composition for treating inflammatory diseases e.g. arthritis.

XX

PS Example 2; Page 107; 196pp; English.

XX

CC The invention relates to immunogenic analogues of multimeric proteins  
CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis  
CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic  
CC analogues. The immunogenic analogue is useful for preparing a composition  
CC for treating inflammatory diseases, e.g., arthritis. It is also used in  
CC gene therapy. The present sequence is a pan DR epitope peptide (PADRE).  
CC This sequence is used to illustrate the method of the invention

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 6; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13

|||||||

Db 1 AKFVAAWTLKAAA 13

#### RESULT 13

AAO30463

ID AAO30463 standard; peptide; 13 AA.

XX

AC AAO30463;

XX

DT 22-SEP-2003 (first entry)

XX

DE Pan DR epitope peptide (PADRE) #2.

XX

KW Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;  
KW tumour necrosis factor alpha; gene therapy; arthritis; epitope; PADRE;  
KW pan DR epitope peptide.

XX

OS Unidentified.

XX



PN WO2003042244-A2.  
 XX  
 PD 22-MAY-2003.  
 XX  
 PF 15-NOV-2002; 2002WO-DK000764.  
 XX  
 PR 16-NOV-2001; 2001DK-00001702.  
 PR 16-NOV-2001; 2001US-0331575P.  
 XX  
 PA (PHAR-) PHARMEXA AS.  
 PA (KLYS/) KLYSNER S.  
 PA (NIEL/) NIELSEN F S.  
 PA (BRAT/) BRATT T.  
 PA (VOLD/) VOLDBORG B.  
 PA (MOUR/) MOURITSEN S.  
 XX  
 PI Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;  
 XX  
 DR WPI; 2003-449558/42.  
 DR N-PSDB; AAL61298.  
 XX  
 PT New immunogenic analogue of a polymeric protein, useful for preparing a  
 PT composition for treating inflammatory diseases e.g. arthritis.  
 XX  
 PS Example 7; Page 124; 196pp; English.  
 XX  
 CC The invention relates to immunogenic analogues of multimeric proteins  
 CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis  
 CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic  
 CC analogues. The immunogenic analogue is useful for preparing a composition  
 CC for treating inflammatory diseases, e.g., arthritis. It is also used in  
 CC gene therapy. The present sequence is a pan DR epitope peptide (PADRE).  
 CC This sequence is used to illustrate the method of the invention  
 XX  
 SQ Sequence 13 AA;  
  
 Query Match 100.0%; Score 64; DB 6; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 AKFVAAWTLKAAA 13  
 | | | | | | | | | |  
 Db 1 AKFVAAWTLKAAA 13

# RESULT 14

ABR82484

ID ABR82484 standard; peptide; 13 AA.

XX

AC ABR82484;

XX

DT 20-NOV-2003 (first entry)

XX

DE Universal MHC class II DR binding sequence.

XX

KW CEA; immune response; carcinoembryonic antigen; antigen presenting cell;

KW APC; cytostatic; vaccine; MHC.

XX  
 OS Synthetic.  
 XX  
 PN WO2003059379-A2.  
 XX  
 PD 24-JUL-2003.  
 XX  
 PF 17-JAN-2003; 2003WO-DK000031.  
 XX  
 PR 17-JAN-2002; 2002DK-00000082.  
 PR 17-JAN-2002; 2002US-0350047P.  
 XX  
 PA (PHAR-) PHARMEXA AS.  
 XX  
 PI Klysner S, Voldborg B;  
 XX  
 DR WPI; 2003-587260/55.  
 XX  
 PT Inducing an immune response in humans against autologous carcinoembryonic  
 PT antigen (CEA) comprises administering a modified CEA polypeptide, a  
 PT nucleic acid encoding the polypeptide, or a microorganism expressing the  
 PT polypeptide.  
 XX  
 PS Disclosure; Page 140; 140pp; English.  
 XX  
 CC The invention relates to inducing an immune response against autologous  
 CC carcinoembryonic antigen (CEA) in an animal, e.g. human. The method  
 CC involves effecting uptake and processing by antigen presenting cells  
 CC (APCs) in the animal of at least 1 modified CEA polypeptide or of a  
 CC nucleic acid encoding the modified CEA polypeptide or of a microorganism  
 CC or virus expressing the modified CEA polypeptide to induce a CTL response  
 CC and an antibody response that targets the autologous CEA. The method is  
 CC useful in immunizing actively against diseases characterized by cells  
 CC that express CEA. The present sequence represents an universal MHC class  
 CC II DR binding sequence  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 7; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

# RESULT 15

ADC81611

ID ADC81611 standard; peptide; 13 AA.

XX

AC ADC81611;

XX

DT 01-JAN-2004 (first entry)

XX

DE Universal MHC class II DR binding sequence PADRE epitope SEQ ID NO:4.

XX

KW pain reduction; nociceptive; nociceptor; immune response;  
 KW tumour necrosis factor alpha; TNFalpha; analgesic; vaccine; pain;  
 KW neuropathic pain; epitope.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003075951-A2.  
 XX  
 PD 18-SEP-2003.  
 XX  
 PF 11-MAR-2003; 2003WO-DK000147.  
 XX  
 PR 11-MAR-2002; 2002DK-00000368.  
 PR 11-MAR-2002; 2002US-0363128P.  
 XX  
 PA (PHAR-) PHARMEXA AS.  
 XX  
 PI Pedersen HR, Ebert B, Pedersen LH, Rasmussen PB;  
 XX  
 DR WPI; 2003-748335/70.  
 XX  
 PT Reducing pain or increasing the threshold for nociception in an  
 PT individual comprises administering an agent capable of inducing an active  
 PT immune response that targets the individual's autologous tumor necrosis  
 PT factor alpha.  
 XX  
 PS Disclosure; SEQ ID NO 4; 120pp; English.  
 XX  
 CC The present invention describes a method for reducing pain or increasing  
 CC the threshold for nociception in an individual comprising administering  
 CC an agent capable of inducing an active immune response that targets the  
 CC individual's autologous tumour necrosis factor alpha (TNFalpha). The  
 CC agent has analgesic activity, and can be used in a vaccine against  
 CC autologous TNFalpha. The method is useful in reducing pain or increasing  
 CC the threshold for nociception in an individual. The method is especially  
 CC intended for reducing neuropathic pain. The present sequence represents a  
 CC universal MHC class II DR binding sequence PADRE epitope, which is given  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 64; DB 7; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
 |||||  
 Db 1 AKFVAAWTLKAAA 13

Search completed: August 13, 2004, 09:12:42  
 Job time : 3.93651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2004, 09:09:23 ; Search time 0.888889 Seconds  
(without alignments)  
755.030 Million cell updates/sec

Title: US-09-785-215-19  
Perfect score: 64  
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

		8					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
1	64	100.0	13	1	US-08-305-871A-23	Sequence 23, Appl	
2	64	100.0	13	4	US-08-788-822A-27	Sequence 27, Appl	
3	64	100.0	13	4	US-09-543-608A-27	Sequence 27, Appl	
4	64	100.0	80	4	US-09-311-784A-22	Sequence 22, Appl	
5	64	100.0	118	4	US-09-311-784A-8	Sequence 8, Appli	
6	64	100.0	119	4	US-09-311-784A-10	Sequence 10, Appl	
7	64	100.0	122	4	US-09-311-784A-39	Sequence 39, Appl	
8	64	100.0	132	4	US-09-311-784A-37	Sequence 37, Appl	
9	64	100.0	136	4	US-09-311-784A-12	Sequence 12, Appl	
10	64	100.0	138	4	US-09-311-784A-4	Sequence 4, Appli	
11	64	100.0	152	4	US-09-311-784A-20	Sequence 20, Appl	

12	64	100.0	164	4	US-09-311-784A-18	Sequence 18, Appl
13	64	100.0	215	4	US-09-311-784A-2	Sequence 2, Appli
14	64	100.0	249	4	US-09-311-784A-16	Sequence 16, Appl
15	64	100.0	253	4	US-09-311-784A-6	Sequence 6, Appli
16	64	100.0	266	4	US-09-311-784A-14	Sequence 14, Appl
17	61	95.3	13	4	US-09-543-608A-28	Sequence 28, Appl
18	57	89.1	13	1	US-08-305-871A-22	Sequence 22, Appl
19	57	89.1	13	4	US-08-788-822A-26	Sequence 26, Appl
20	57	89.1	13	4	US-09-692-170C-29	Sequence 29, Appl
21	57	89.1	13	4	US-09-543-608A-26	Sequence 26, Appl
22	57	89.1	22	4	US-09-692-170C-17	Sequence 17, Appl
23	57	89.1	22	4	US-09-692-170C-18	Sequence 18, Appl
24	57	89.1	22	4	US-09-692-170C-20	Sequence 20, Appl
25	57	89.1	22	4	US-09-692-170C-21	Sequence 21, Appl
26	57	89.1	22	4	US-09-692-170C-22	Sequence 22, Appl
27	57	89.1	22	4	US-09-692-170C-23	Sequence 23, Appl
28	57	89.1	22	4	US-09-692-170C-24	Sequence 24, Appl
29	57	89.1	23	4	US-09-692-170C-15	Sequence 15, Appl
30	57	89.1	23	4	US-09-692-170C-16	Sequence 16, Appl
31	57	89.1	24	4	US-09-983-019-10	Sequence 10, Appl
32	57	89.1	25	3	US-09-075-257A-17	Sequence 17, Appl
33	57	89.1	25	3	US-09-534-639-17	Sequence 17, Appl
34	57	89.1	25	4	US-09-692-170C-25	Sequence 25, Appl
35	57	89.1	25	4	US-09-692-170C-37	Sequence 37, Appl
36	57	89.1	25	4	US-09-692-170C-44	Sequence 44, Appl
37	57	89.1	27	4	US-09-983-019-4	Sequence 4, Appli
38	57	89.1	27	4	US-09-983-019-7	Sequence 7, Appli
39	57	89.1	28	3	US-09-075-257A-15	Sequence 15, Appl
40	57	89.1	28	3	US-09-075-257A-16	Sequence 16, Appl
41	57	89.1	28	3	US-09-534-639-15	Sequence 15, Appl
42	57	89.1	28	3	US-09-534-639-16	Sequence 16, Appl
43	57	89.1	28	4	US-09-692-170C-35	Sequence 35, Appl
44	57	89.1	28	4	US-09-692-170C-36	Sequence 36, Appl
45	55	85.9	13	4	US-09-543-608A-29	Sequence 29, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-305-871A-23

; Sequence 23, Application US/08305871A

; Patent No. 5736142

#### ; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Gaeta, Federico

; APPLICANT: Grey, Howard M.

; APPLICANT: Sidney, John

; APPLICANT: Alexander, Jeffrey L.

; TITLE OF INVENTION: Alteration of Immune Response Using Pan

; TITLE OF INVENTION: DR-Binding Peptides

; NUMBER OF SEQUENCES: 29

#### ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

```

; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,871A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,101
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-0062-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..13
; OTHER INFORMATION: /note= "Peptide wherein X is
; OTHER INFORMATION: tyrosine or phenylalanine."
US-08-305-871A-23

```

```

Query Match          100.0%; Score 64; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13

```

# RESULT 2

US-08-788-822A-27

; Sequence 27, Application US/08788822A

; Patent No. 6413935

## ; GENERAL INFORMATION:

; APPLICANT: Alexander, Jeffrey L.

; APPLICANT: DeFrees, Shawn

; APPLICANT: Sette, Alessandro

; TITLE OF INVENTION: Induction of Immune Response Against

; TITLE OF INVENTION: Desired Determinants

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

```

; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,822A
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,510
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 014137-009210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-788-822A-27

```

```

Query Match          100.0%; Score 64; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13

```

# RESULT 3

US-09-543-608A-27

```

; Sequence 27, Application US/09543608A
; Patent No. 6602510
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Celis, Esteban
; APPLICANT: Keogh, Elissa A.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: HLA Class I A2 Tumor Associated Antigen
; TITLE OF INVENTION: Peptides and Vaccine Compositions

```

; FILE REFERENCE: 018623-015710US  
; CURRENT APPLICATION NUMBER: US/09/543,608A  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PanDR binding peptide (PADRE)  
US-09-543-608A-27

Query Match 100.0%; Score 64; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 4

US-09-311-784A-22  
; Sequence 22, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SigTh  
US-09-311-784A-22

Query Match 100.0%; Score 64; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |



Db 68 AKFVAAWTLKAAA 80

RESULT 5

US-09-311-784A-8

; Sequence 8, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: KappaLAMP-Th  
US-09-311-784A-8

Query Match 100.0%; Score 64; DB 4; Length 118;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 68 AKFVAAWTLKAAA 80

RESULT 6

US-09-311-784A-10

; Sequence 10, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A

; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: H2M-Th  
US-09-311-784A-10

Query Match 100.0%; Score 64; DB 4; Length 119;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 63 AKFVAAWTLKAAA 75

RESULT 7

US-09-311-784A-39  
; Sequence 39, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: vector pMIN.1  
US-09-311-784A-39

Query Match 100.0%; Score 64; DB 4; Length 122;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |

Db 34 AKFVAAWTLKAAA 46

RESULT 8

US-09-311-784A-37  
; Sequence 37, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 132  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: vector pMIN.0  
US-09-311-784A-37

Query Match 100.0%; Score 64; DB 4; Length 132;  
Best Local Similarity 100.0%; Pred. No. 0.0025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 34 AKFVAAWTLKAAA 46

RESULT 9

US-09-311-784A-12  
; Sequence 12, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A

; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: H2O-Th  
US-09-311-784A-12

Query Match 100.0%; Score 64; DB 4; Length 136;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 67 AKFVAAWTLKAAA 79

RESULT 10

US-09-311-784A-4  
; Sequence 4, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: I80T  
US-09-311-784A-4

Query Match 100.0%; Score 64; DB 4; Length 138;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |

Db 126 AKFVAAWTLKAAA 138

RESULT 11

US-09-311-784A-20

; Sequence 20, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Ig-betaTh  
US-09-311-784A-20

Query Match 100.0%; Score 64; DB 4; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 71 AKFVAAWTLKAAA 83

RESULT 12

US-09-311-784A-18

; Sequence 18, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A

; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Ig-alphaTh  
US-09-311-784A-18

Query Match 100.0%; Score 64; DB 4; Length 164;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
|||||||  
Db 74 AKFVAAWTLKAAA 86

RESULT 13

US-09-311-784A-2  
; Sequence 2, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IiPADRE  
US-09-311-784A-2

Query Match 100.0%; Score 64; DB 4; Length 215;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
|||||||

Db

88 AKFVAAWTLKAAA 100

RESULT 14

US-09-311-784A-16

; Sequence 16, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PADRE-HBV-s  
US-09-311-784A-16

Query Match 100.0%; Score 64; DB 4; Length 249;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 2 AKFVAAWTLKAAA 14

RESULT 15

US-09-311-784A-6

; Sequence 6, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A

; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IiThfull  
US-09-311-784A-6

Query Match 100.0%; Score 64; DB 4; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
|||||||  
Db 126 AKFVAAWTLKAAA 138

Search completed: August 13, 2004, 09:17:36  
Job time : 0.888889 secs



OM protein - protein search, using sw model

Run on: August 13, 2004, 09:08:23 ; Search time 0.714286 Seconds  
(without alignments)  
1750.686 Million cell updates/sec

Title: US-09-785-215-19  
Perfect score: 64  
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		8				
Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	39	60.9	146	2	D70766	hypothetical prote
2	39	60.9	261	2	AH0510	carnitine racemase
3	39	60.9	297	1	D64724	carnitine racemase
4	39	60.9	297	2	G85484	carnitine racemase
5	39	60.9	297	2	G90633	carnitine racemase
6	38	59.4	155	2	A29699	beta-lactoglobulin
7	38	59.4	204	2	A53848	ubiquitin-conjugat
8	38	59.4	352	2	AE3610	probable transmemb
9	38	59.4	411	2	S61245	probable virion gl
10	38	59.4	639	2	D95884	hypothetical prote
11	38	59.4	758	2	B82122	ferrous iron trans
12	37	57.8	309	2	F70796	hypothetical prote
13	37	57.8	368	2	G87458	conserved hypothet

14	37	57.8	410	2	E75190	probable 2,3-bisph
15	37	57.8	712	2	AG1888	Na <sup>+</sup> /H <sup>+</sup> antiporter
16	36	56.2	156	2	T02806	probable membrane
17	36	56.2	317	2	F87336	conserved hypothet
18	36	56.2	330	2	G89771	lipoprotein [impor
19	36	56.2	357	2	T12379	NADH2 dehydrogenas
20	36	56.2	362	2	C90674	hypothetical prote
21	36	56.2	362	2	C64758	yahA protein - Esc
22	36	56.2	363	2	B82532	GTP-binding protei
23	36	56.2	365	2	F85524	hypothetical prote
24	36	56.2	412	2	T22076	hypothetical prote
25	36	56.2	569	2	T00851	hypothetical prote
26	36	56.2	690	2	T23399	hypothetical prote
27	36	56.2	776	1	JH0570	hypothetical prote
28	36	56.2	1031	2	D88912	transferrin recept
29	36	56.2	1031	2	T33655	protein T06A10.1 [
30	36	56.2	2076	2	T28915	hypothetical prote
31	35	54.7	120	2	AF0299	probable membrane
32	35	54.7	128	2	C48552	hypothetical prote
33	35	54.7	226	2	G81894	probable integral
34	35	54.7	226	2	D81125	hypothetical prote
35	35	54.7	301	2	B90027	hypothetical prote
36	35	54.7	388	2	AF2718	ribonuclease D [im
37	35	54.7	429	2	B97500	ribonuclease D (PA
38	35	54.7	550	2	A57519	parathyroid hormon
39	35	54.7	587	2	T01829	hypothetical prote
40	35	54.7	639	2	S03547	hypothetical prote
41	35	54.7	659	2	A72729	hypothetical prote
42	35	54.7	723	2	T14605	probable cell divi
43	35	54.7	730	2	T44246	translation elonga
44	35	54.7	750	2	A56881	prostate-specific
45	35	54.7	773	2	A82998	probable oxidoredu

#### ALIGNMENTS

##### RESULT 1

D70766

hypothetical protein Rv2081c - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C;Accession: D70766

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekalia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: D70766

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
A;Residues: 1-146 <COL>  
A;Cross-references: GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98193.1;  
PID:e247091; PID:g1370247  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: Rv2081c

Query Match 60.9%; Score 39; DB 2; Length 146;  
Best Local Similarity 63.6%; Pred. No. 8.2;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13  
| | | | : | |  
Db 10 FVAIWTARAAS 20

RESULT 2

AH0510

carnitine racemase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AH0510

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Sebaihia, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Feltwell, T.; Hamlin, N.; Haque, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AH0510

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-261 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD01224.1; PID:g16501353; GSPDB:GN00176

C;Genetics:

A;Gene: STY0080

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 261;  
Best Local Similarity 72.7%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13  
| | | | | | |  
Db 60 FSAGWDLKAAA 70

RESULT 3

D64724

carnitine racemase (EC 5.-.-.-) - Escherichia coli (strain K-12)  
 C;Species: Escherichia coli  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
 C;Accession: D64724; S40557; I41014  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;  
 Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor,  
 J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of Escherichia coli K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: D64724  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-297 <BLAT>  
 A;Cross-references: GB:AE000114; GB:U00096; NID:g1786217; PIDN:AAC73147.1;  
 PID:g1786220; UWGP:b0036  
 A;Experimental source: strain K-12, substrain MGL655  
 R;Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono,  
 K.; Mizobuchi, K.; Nakata, A.  
 submitted to the EMBL Data Library, December 1992  
 A;Description: Systematic sequencing of the Escherichia coli genome: analysis of  
 the 0-2.4min region.  
 A;Reference number: S40531  
 A;Accession: S40557  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-144,'A',146-297 <YUR>  
 A;Cross-references: EMBL:D10483; NID:g216434; PIDN:BAA01312.1; PID:g216461  
 R;Eichler, K.; Bourgis, F.; Buchet, A.; Kleber, H.P.; Mandrand-Berthelot, M.A.  
 Mol. Microbiol. 13, 775-786, 1994  
 A;Title: Molecular characterization of the cai operon necessary for carnitine  
 metabolism in Escherichia coli.  
 A;Reference number: I41010; MUID:95115548; PMID:7815937  
 A;Accession: I41014  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-2,'R',4-13,'A',15-282,'L',284-297 <RES>  
 A;Cross-references: EMBL:X73904; NID:g563860; PIDN:CAA52114.1; PID:g563865  
 A;Experimental source: strain 044 K74  
 C;Genetics:  
 A;Gene: caiD  
 C;Function:  
 A;Pathway: carnitine metabolism  
 C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology  
 C;Keywords: isomerase  
 F;60-212/Domain: enoyl-CoA hydratase homology <ECH>

Query Match 60.9%; Score 39; DB 1; Length 297;  
 Best Local Similarity 72.7%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 96 FSAGWDLKAAA 106

RESULT 4

G85484

carnitine racemase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002

C;Accession: G85484

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.  
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G85484

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-297 <STO>

A;Cross-references: GB:AE005174; NID:g12512723; PIDN:AAG54339.1; GSPDB:GN00145; UWGP:Z0042

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: caid

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 297;

Best Local Similarity 72.7%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13

| | | | |

Db 96 FSAGWDLKAAA 106

RESULT 5

G90633

carnitine racemase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 17-May-2002

C;Accession: G90633

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: G90633

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-297 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA33462.1; PID:g13359495; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs0039

C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 60.9%; Score 39; DB 2; Length 297;  
Best Local Similarity 72.7%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13  
| | | | |  
Db 96 FSAGWDLKAAA 106

RESULT 6

A29699

beta-lactoglobulin - eastern gray kangaroo

C;Species: Macropus giganteus (eastern gray kangaroo)

C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 18-Jul-1997

C;Accession: A29699

R;Godovac-Zimmermann, J.; Shaw, D.

Biol. Chem. Hoppe-Seyler 368, 879-886, 1987

A;Title: The primary structure, binding site and possible function of beta-lactoglobulin from eastern grey kangaroo (Macropus giganteus).

A;Reference number: A29699; MUID:87299024; PMID:3620116

A;Accession: A29699

A;Molecule type: protein

A;Residues: 1-155 <GOD>

C;Superfamily: lipocalin; lipocalin homology

F;10-155/Domain: lipocalin homology <LIP>

Query Match 59.4%; Score 38; DB 2; Length 155;  
Best Local Similarity 58.3%; Pred. No. 13;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13  
| | | : | | : | |  
Db 14 KFGVSWYLREAA 25

RESULT 7

A53848

ubiquitin-conjugating enzyme - yeast (Pichia pastoris)

C;Species: Pichia pastoris

C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 17-Mar-2000

C;Accession: A53848

R;Crane, D.I.; Kalish, J.E.; Gould, S.J.

J. Biol. Chem. 269, 21835-21844, 1994

A;Title: The Pichia pastoris PAS4 gene encodes a ubiquitin-conjugating enzyme required for peroxisome assembly.

A;Reference number: A53848; MUID:94342381; PMID:8063827

A;Accession: A53848

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-204 <CRA>

A;Cross-references: GB:U12511; NID:g531386; PIDN:AAA53634.1; PID:g531387

C;Genetics:

A;Gene: PAS4

C;Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match 59.4%; Score 38; DB 2; Length 204;  
Best Local Similarity 58.3%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12  
||: |||| :|  
Db 139 AKWTPAWTLSSA 150

RESULT 8

AE3610

probable transmembrane protein [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002

C;Accession: AE3610

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.; Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn, R.; Kyrpides, N.; Overbeek, R.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.

A;Reference number: AD3252; PMID:11756688

A;Accession: AE3610

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-352 <KUR>

A;Cross-references: GB:AE008918; PIDN:AAL54048.1; PID:g17985002; GSPDB:GN00191

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEII0806

A;Map position: II

Query Match 59.4%; Score 38; DB 2; Length 352;  
Best Local Similarity 63.6%; Pred. No. 29;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11  
| |:|||:| |  
Db 331 AMFIAAWSLLA 341

RESULT 9

S61245

probable virion glycoprotein M (gM) - bovine herpesvirus 1

C;Species: bovine herpesvirus 1

C;Date: 18-Sep-1997 #sequence\_revision 18-Sep-1997 #text\_change 26-Aug-1999

C;Accession: S61245

R;Vlcek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Schwzyer, M.

submitted to the EMBL Data Library, January 1995

A;Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of herpes simplex virus.

A;Reference number: S61233

A;Accession: S61245

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-411 <VLC>  
A;Cross-references: EMBL:Z48053; NID:g971311; PIDN:CAA88123.1; PID:g971324  
C;Superfamily: herpesvirus 51K protein

Query Match 59.4%; Score 38; DB 2; Length 411;  
Best Local Similarity 77.8%; Pred. No. 34;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VAAWTLKAA 12  
:||||:|  
Db 141 LAAWTLQAA 149

RESULT 10

D95884

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021)  
magaplasmid pSymB

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C;Accession: D95884

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez-Lucas, I.; Becker, A.; Cowie, A.; Gouzy, J.; Golding, B.; Puhler, A.

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: D95884

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-639 <KUR>

A;Cross-references: GB:AL591985; PIDN:CAC48740.1; PID:g15140213; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; Bothe, G.; Boutry, M.; Bowser, L.; Buhrmester, J.; Cadieu, E.; Capela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Gloux, S.; Godrie, T.; Goffeau, A.; Golding, B.; Gouzy, J.; Gurjal, M.; Hernandez-Lucas, I.; Hong, A.; Huizar, L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, V.; Masuy, D.; Palm, C.; Peck, M.C.; Pohl, T.M.; Portetelle, D.; Purnelle, B.; Ramsperger, U.; Surzycki, R.; Thebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.; Batut, J.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMb20354

A;Genome: plasmid

Query Match 59.4%; Score 38; DB 2; Length 639;  
Best Local Similarity 69.2%; Pred. No. 52;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;



QY 1 AKFVAAWTLKAAA 13  
| | | | | | |  
Db 491 ALFGAAWTLAYAA 503

RESULT 11

B82122

ferrous iron transport protein B VC2077 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 24-Aug-2001

C;Accession: B82122

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82122

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-758 <HEI>

A;Cross-references: GB:AE004281; GB:AE003852; NID:g9656616; PIDN:AAF95223.1; GSPDB:GN00126; TIGR:VC2077

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2077

A;Map position: 1

C;Superfamily: ferrous iron transport protein B; translation elongation factor Tu homology

Query Match 59.4%; Score 38; DB 2; Length 758;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKFVAAWTLKAA 12  
| : | : | : |  
Db 694 ARFIAVWTMGLA 705

RESULT 12

F70796

hypothetical protein Rv3725 - *Mycobacterium tuberculosis* (strain H37RV)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C;Accession: F70796

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaiia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: F70796

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-309 <COL>

A;Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA18047.1; PID:g2960149

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: Rv3725

C;Superfamily: dihydrokaempferol 4-reductase

Query Match 57.8%; Score 37; DB 2; Length 309;  
Best Local Similarity 63.6%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
|| || || |  
Db 16 FVGWGTAKAIA 26

#### RESULT 13

G87458

conserved hypothetical protein CC1690 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: G87458

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: G87458

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-368 <STO>

A;Cross-references: GB:AE005673; NID:g13423101; PIDN:AAK23667.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC1690

Query Match 57.8%; Score 37; DB 2; Length 368;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKFVAAWT 8  
| |||||  
Db 342 APFVAAWT 349

RESULT 14

E75190

probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1) - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 27-Oct-2003

C;Accession: E75190

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution.

A;Reference number: A75001

A;Accession: E75190

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-410 <KAW>

A;Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB48972.1; PID:g5457481

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2318

C;Superfamily: cofactor-independent phosphoglycerate mutase, archaeal-type

Query Match 57.8%; Score 37; DB 2; Length 410;

Best Local Similarity 58.3%; Pred. No. 51;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13

|| | :|||

Db 235 KFTEQWKVKAAA 246

RESULT 15

AG1888

Na<sup>+</sup>/H<sup>+</sup> antiporter [imported] - *Nostoc* sp. (strain PCC 7120)

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AG1888

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.;

Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.;

Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.;

Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AG1888

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-712 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA072614.1; PID:g17130002; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr0656

Query Match 57.8%; Score 37; DB 2; Length 712;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKFVAAWTLK 10  
:|:|:| |  
Db 323 SKFIAAWLAK 332

Search completed: August 13, 2004, 09:16:34  
Job time : 2.71429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2004, 09:15:49 ; Search time 2.44444 Seconds  
(without alignments)  
1669.522 Million cell updates/sec

Title: US-09-785-215-19  
Perfect score: 64  
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result  
No.      Score      Match Length DB      ID      Description

1	64	100.0	13	9	US-09-894-018-69	Sequence 69, Appl
2	64	100.0	13	9	US-09-949-375A-18	Sequence 18, Appl
3	64	100.0	13	9	US-09-785-215-19	Sequence 19, Appl
4	64	100.0	13	14	US-10-223-809A-17	Sequence 17, Appl
5	64	100.0	13	14	US-10-295-074-7	Sequence 7, Appli
6	64	100.0	13	14	US-10-295-074-20	Sequence 20, Appl
7	64	100.0	13	15	US-10-371-069-52	Sequence 52, Appl
8	64	100.0	13	15	US-10-371-645-52	Sequence 52, Appl
9	64	100.0	13	15	US-10-371-260-52	Sequence 52, Appl
10	64	100.0	13	16	US-10-441-779C-33	Sequence 33, Appl
11	64	100.0	55	9	US-09-949-375A-12	Sequence 12, Appl
12	64	100.0	80	9	US-09-894-018-97	Sequence 97, Appl
13	64	100.0	80	12	US-10-371-525-22	Sequence 22, Appl
14	64	100.0	80	15	US-10-371-069-22	Sequence 22, Appl
15	64	100.0	80	15	US-10-371-645-22	Sequence 22, Appl
16	64	100.0	80	15	US-10-371-260-22	Sequence 22, Appl
17	64	100.0	98	9	US-09-894-018-101	Sequence 101, App
18	64	100.0	106	9	US-09-894-018-95	Sequence 95, Appl
19	64	100.0	107	9	US-09-894-018-93	Sequence 93, Appl
20	64	100.0	107	9	US-09-894-018-103	Sequence 103, App
21	64	100.0	118	12	US-10-371-525-8	Sequence 8, Appli
22	64	100.0	118	15	US-10-371-069-8	Sequence 8, Appli
23	64	100.0	118	15	US-10-371-645-8	Sequence 8, Appli
24	64	100.0	118	15	US-10-371-260-8	Sequence 8, Appli
25	64	100.0	119	12	US-10-371-525-10	Sequence 10, Appl
26	64	100.0	119	15	US-10-371-069-10	Sequence 10, Appl
27	64	100.0	119	15	US-10-371-645-10	Sequence 10, Appl
28	64	100.0	119	15	US-10-371-260-10	Sequence 10, Appl
29	64	100.0	122	12	US-10-371-525-39	Sequence 39, Appl
30	64	100.0	122	15	US-10-371-069-39	Sequence 39, Appl
31	64	100.0	122	15	US-10-371-645-39	Sequence 39, Appl
32	64	100.0	122	15	US-10-371-260-39	Sequence 39, Appl
33	64	100.0	123	9	US-09-894-018-109	Sequence 109, App
34	64	100.0	130	9	US-09-894-018-99	Sequence 99, Appl
35	64	100.0	132	12	US-10-371-525-37	Sequence 37, Appl
36	64	100.0	132	15	US-10-371-069-37	Sequence 37, Appl
37	64	100.0	132	15	US-10-371-645-37	Sequence 37, Appl
38	64	100.0	132	15	US-10-371-260-37	Sequence 37, Appl
39	64	100.0	136	12	US-10-371-525-12	Sequence 12, Appl
40	64	100.0	136	15	US-10-371-069-12	Sequence 12, Appl
41	64	100.0	136	15	US-10-371-645-12	Sequence 12, Appl
42	64	100.0	136	15	US-10-371-260-12	Sequence 12, Appl
43	64	100.0	138	12	US-10-371-525-4	Sequence 4, Appli
44	64	100.0	138	15	US-10-371-069-4	Sequence 4, Appli
45	64	100.0	138	15	US-10-371-645-4	Sequence 4, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-894-018-69

; Sequence 69, Application US/09894018

; Patent No. US20020119127A1

; GENERAL INFORMATION:

; APPLICANT: EPIMMUNE, Inc.

```

; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE sequence
US-09-894-018-69

```

```

Query Match          100.0%; Score 64; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13

```

```

RESULT 2
US-09-949-375A-18
; Sequence 18, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 17.
US-09-949-375A-18

```

```

Query Match          100.0%; Score 64; DB 9; Length 13;

```

Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
|||||||  
Db 1 AKFVAAWTLKAAA 13

RESULT 3

US-09-785-215-19

; Sequence 19, Application US/09785215

; Publication No. US20020187157A1

; GENERAL INFORMATION:

; APPLICANT: JENSEN, Martin Roland et al.

; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID

; FILE REFERENCE: 3631-0107P

; CURRENT APPLICATION NUMBER: US/09/785,215

; CURRENT FILING DATE: 2001-02-20

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 19

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: artificial T-cell epitope capable of binding to a large portion

; OTHER INFORMATION: of MHC Class II molecules in a variety of animals

US-09-785-215-19

Query Match 100.0%; Score 64; DB 9; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00057;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
|||||||  
Db 1 AKFVAAWTLKAAA 13

RESULT 4

US-10-223-809A-17

; Sequence 17, Application US/10223809A

; Publication No. US20030157117A1

; GENERAL INFORMATION:

; APPLICANT: Pharmexa A/S

; APPLICANT: Rasmussen, Peter Birk et al.

; TITLE OF INVENTION: No. US20030157117A1 Method for Down-Regulation of Amyloid

; FILE REFERENCE: 674542-2008

; CURRENT APPLICATION NUMBER: US/10/223,809A

; CURRENT FILING DATE: 2002-08-20

; PRIOR APPLICATION NUMBER: US 60/337,543

; PRIOR FILING DATE: 2001-08-20

; PRIOR APPLICATION NUMBER: US 60/373,027

; PRIOR FILING DATE: 2002-04-16

; PRIOR APPLICATION NUMBER: DE 2001 01231

; PRIOR FILING DATE: 2001-08-20



; PRIOR APPLICATION NUMBER: DE 2002 0058  
; PRIOR FILING DATE: 2002-04-16  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HLA DR binding sequence  
US-10-223-809A-17

Query Match 100.0%; Score 64; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 5

US-10-295-074-7  
; Sequence 7, Application US/10295074  
; Publication No. US20030185845A1  
; GENERAL INFORMATION:  
; APPLICANT: Pharmexa A/S  
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS  
; FILE REFERENCE: P1013DK00  
; CURRENT APPLICATION NUMBER: US/10/295,074  
; CURRENT FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Pan DR binding peptide (PADRE)  
US-10-295-074-7

Query Match 100.0%; Score 64; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 6

US-10-295-074-20  
; Sequence 20, Application US/10295074  
; Publication No. US20030185845A1  
; GENERAL INFORMATION:  
; APPLICANT: Pharmexa A/S  
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS

; FILE REFERENCE: P1013DK00  
; CURRENT APPLICATION NUMBER: US/10/295,074  
; CURRENT FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Pan DR binding peptide (PADRE)  
US-10-295-074-20

Query Match 100.0%; Score 64; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 7

US-10-371-069-52  
; Sequence 52, Application US/10371069  
; Publication No. US20030216342A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMMUNE Inc.  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.10  
; CURRENT APPLICATION NUMBER: US/10/371,069  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/078,904  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 52  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PADRE-universal MHC class II epitope  
US-10-371-069-52

Query Match 100.0%; Score 64; DB 15; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 8

US-10-371-645-52  
; Sequence 52, Application US/10371645  
; Publication No. US20030216343A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMMUNE Inc.  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.11  
; CURRENT APPLICATION NUMBER: US/10/371,645  
; CURRENT FILING DATE: 2003-06-20  
; PRIOR APPLICATION NUMBER: US 09/078,904  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 52  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PADRE-universal MHC class II epitope  
US-10-371-645-52

Query Match 100.0%; Score 64; DB 15; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 9

US-10-371-260-52  
; Sequence 52, Application US/10371260  
; Publication No. US20030220285A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMMUNE Inc.  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.

```
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PADRE-universal MHC class II epitope
US-10-371-260-52
```

```
Query Match          100.0%; Score 64; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      1 AKFVAAWTLKAAA 13
```

RESULT 10

US-10-441-779C-33

```
; Sequence 33, Application US/10441779C
; Publication No. US20040141958A1
; GENERAL INFORMATION:
; APPLICANT: Steinaa, Lucilla
; APPLICANT: Mouritsen, Soren
; APPLICANT: Gautam, Anand
; APPLICANT: Haaning, Jesper
; APPLICANT: Dalum, Iben
; APPLICANT: Birk, Peter
; APPLICANT: Leach, Dana
; APPLICANT: Nielsen, Klauss
; APPLICANT: Karlsson, Gunilla
; TITLE OF INVENTION: NOVEL METHODS FOR THERAPEUTIC VACCINATION
; FILE REFERENCE: 4614-0116P
; CURRENT APPLICATION NUMBER: US/10/441,779C
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 09/413,186
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/105,011
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: PA 1998 01261
; PRIOR FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
```

; SEQ ID NO 33  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pan DR epitope peptide  
US-10-441-779C-33

Query Match 100.0%; Score 64; DB 16; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 1 AKFVAAWTLKAAA 13

RESULT 11

US-09-949-375A-12  
; Sequence 12, Application US/09949375A  
; Patent No. US20020172673A1  
; GENERAL INFORMATION:  
; APPLICANT: KLYSNER, Steen et al.  
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE  
; FILE REFERENCE: 3631-0111P  
; CURRENT APPLICATION NUMBER: US/09/949,375A  
; CURRENT FILING DATE: 2002-01-18  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic amino acid sequence of SEQ ID NO: 11.  
US-09-949-375A-12

Query Match 100.0%; Score 64; DB 9; Length 55;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 13 AKFVAAWTLKAAA 25

RESULT 12

US-09-894-018-97  
; Sequence 97, Application US/09894018  
; Patent No. US20020119127A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Chestnut, Robert  
; APPLICANT: Livingston, Brian  
; APPLICANT: Baker, Denisw  
; APPLICANT: Newman, Mark

```

; APPLICANT: Brown, David
; TITLE OF INVENTION: METODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s2(-3)
US-09-894-018-97

```

```

Query Match          100.0%; Score 64; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      32 AKFVAAWTLKAAA 44

```

# RESULT 13

US-10-371-525-22

```

; Sequence 22, Application US/10371525
; Publication No. US20030203869A1
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/10/371,525
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/311,784
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT

```

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SigTh  
US-10-371-525-22

Query Match 100.0%; Score 64; DB 12; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0033;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 68 AKFVAAWTLKAAA 80

RESULT 14

US-10-371-069-22

; Sequence 22, Application US/10371069  
; Publication No. US20030216342A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMMUNE Inc.  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.10  
; CURRENT APPLICATION NUMBER: US/10/371,069  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/078,904  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SigTh  
US-10-371-069-22

Query Match 100.0%; Score 64; DB 15; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0033;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAAA 13  
| | | | | | | | | |  
Db 68 AKFVAAWTLKAAA 80

RESULT 15

US-10-371-645-22

```

; Sequence 22, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SigTh
US-10-371-645-22

```

```

Query Match          100.0%; Score 64; DB 15; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AKFVAAWTLKAAA 13
        |||||
Db      68 AKFVAAWTLKAAA 80

```

```

Search completed: August 13, 2004, 09:33:25
Job time : 3.44444 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2004, 09:07:53 ; Search time 1.98413 Seconds  
(without alignments)  
2067.275 Million cell updates/sec

Title: US-09-785-215-19  
Perfect score: 64  
Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length	DB	ID		Description

---

1	43	67.2	370	16	Q8R8U0	Q8r8u0 thermoanaer
2	42	65.6	2223	2	Q7WTF4	Q7wtf4 streptomyce
3	41	64.1	226	16	Q7U559	Q7u559 synechococc
4	41	64.1	832	16	Q8XYF4	Q8xyf4 ralstonia s
5	40	62.5	275	16	Q98CR0	Q98cr0 rhizobium l
6	40	62.5	422	2	Q9ETG0	Q9etg0 corynebacte
7	39	60.9	147	16	Q7TZ25	Q7tzz25 mycobacteri
8	39	60.9	285	5	Q7YTL3	Q7ytl3 caenorhabdi
9	39	60.9	331	5	Q8STE0	Q8ste0 encephalito
10	39	60.9	612	5	Q8STA9	Q8sta9 encephalito
11	39	60.9	758	16	Q8DFQ8	Q8dfq8 vibrio vuln
12	38	59.4	96	2	Q93DU6	Q93du6 unidentifie
13	38	59.4	352	16	Q8YBT0	Q8ybt0 brucella me
14	38	59.4	352	16	Q8FWJ3	Q8fwj3 brucella su
15	38	59.4	438	12	O39493	O39493 bovine herp
16	38	59.4	518	10	Q7XWU9	Q7xwu9 oryza sativ
17	38	59.4	639	16	Q92WJ9	Q92wj9 rhizobium m
18	38	59.4	698	10	Q94HP0	Q94hp0 oryza sativ
19	38	59.4	698	10	Q7XGK4	Q7xgk4 oryza sativ
20	38	59.4	758	16	Q9KQC3	Q9kqc3 vibrio chol
21	38	59.4	758	16	Q87RD9	Q87rd9 vibrio para
22	38	59.4	826	16	Q8EIJ1	Q8eij1 shewanella
23	38	59.4	916	3	Q875W4	Q875w4 saccharomyc
24	38	59.4	3037	12	Q68749	Q68749 hepatitis c
25	37	57.8	130	16	Q879S9	Q879s9 xylella fas
26	37	57.8	131	2	Q847G5	Q847g5 pseudomonas
27	37	57.8	288	16	Q989E2	Q989e2 rhizobium l
28	37	57.8	309	16	O69692	O69692 mycobacteri
29	37	57.8	329	5	Q9VFY8	Q9vfy8 drosophila
30	37	57.8	333	16	Q7TVU4	Q7tvu4 mycobacteri
31	37	57.8	368	16	Q9A7N1	Q9a7n1 caulobacter
32	37	57.8	391	5	Q9NKE6	Q9nke6 drosophila
33	37	57.8	490	3	Q8J0I9	Q8j0i9 trichoderma
34	37	57.8	612	16	Q987K9	Q987k9 rhizobium l
35	37	57.8	663	3	Q96X20	Q96x20 laccaria fr
36	37	57.8	678	5	Q9I7L8	Q9i7l8 drosophila
37	37	57.8	712	16	Q8YZ31	Q8yz31 anabaena sp
38	37	57.8	783	16	Q7WF63	Q7wf63 bordetella
39	37	57.8	783	16	Q7W3T4	Q7w3t4 bordetella
40	37	57.8	907	5	Q8IF42	Q8if42 trypanosoma
41	37	57.8	1093	3	Q872P9	Q872p9 neurospora
42	37	57.8	1205	3	O59946	O59946 emericella
43	37	57.8	1470	5	Q9VF24	Q9vf24 drosophila
44	37	57.8	3032	12	Q99IB2	Q99ib2 hepatitis c
45	37	57.8	3032	12	Q99IB4	Q99ib4 hepatitis c

#### ALIGNMENTS

##### RESULT 1

Q8R8U0

ID Q8R8U0 PRELIMINARY; PRT; 370 AA.  
AC Q8R8U0;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Kef-type K<sup>+</sup> transport systems, membrane components.  
 GN KEFB OR TTE1901.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL; AE013140; AAM25084.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015299; F:solute:hydrogen antiporter activity; IEA.  
 DR GO; GO:0006885; P:regulation of pH; IEA.  
 DR InterPro; IPR006153; Na\_H\_porter.  
 DR Pfam; PF00999; Na\_H\_Exchanger; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 370 AA; 40695 MW; 489870521F84215A CRC64;

Query Match 67.2%; Score 43; DB 16; Length 370;  
 Best Local Similarity 80.0%; Pred. No. 18;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 FVAAWTLKAA 12  
 ||| |||:||  
 Db 101 FVAGWTLQAA 110

## RESULT 2

### Q7WTF4

ID Q7WTF4 PRELIMINARY; PRT; 2223 AA.  
 AC Q7WTF4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Nana2.  
 GN NANA2.  
 OS Streptomyces nanchangensis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=204925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS3226;  
 RX MEDLINE=22654646; PubMed=12770825;  
 RA Sun Y., Zhou X., Dong H., Tu G., Wang M., Wang B., Deng Z.;  
 RT "A Complete Gene Cluster from Streptomyces nanchangensis NS3226  
 RT Encoding Biosynthesis of the Polyether Ionophore Nanchangmycin.";  
 RL Chem. Biol. 10:431-441(2003).  
 DR EMBL; AF521085; AAP42856.1; -.  
 SQ SEQUENCE 2223 AA; 229037 MW; 6E7ED17FBD94BB3A CRC64;

Query Match 65.6%; Score 42; DB 2; Length 2223;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
: ||||: ||||  
Db 1863 RLVAAWSAKAAA 1874

RESULT 3

Q7U559

ID Q7U559 PRELIMINARY; PRT; 226 AA.  
AC Q7U559;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC  
DE 2.7.7.60).  
GN ISPD OR SYNW1849.  
OS Synechococcus sp. (strain WH8102).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
OX NCBI\_TaxID=84588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825697; PubMed=12917641;  
RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain P.,  
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,  
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;  
RT "The genome of a motile marine Synechococcus."  
RL Nature 424:1037-1042(2003).  
DR EMBL; BX569694; CAE08364.1; -.  
KW Nucleotidyltransferase; Transferase; Complete proteome.  
SQ SEQUENCE 226 AA; 24241 MW; 4B499181FE2EA5A3 CRC64;

Query Match 64.1%; Score 41; DB 16; Length 226;  
Best Local Similarity 81.8%; Pred. No. 24;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAA 12  
| | |||||  
Db 28 KPVIATWLKAA 38

RESULT 4

Q8XYF4

ID Q8XYF4 PRELIMINARY; PRT; 832 AA.  
AC Q8XYF4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein RSC1804.  
GN RSC1804 OR RS04224.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=GMI1000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choisine N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646066; CAD15506.1; -.  
 DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR008278; 4-PPT\_transf.  
 DR InterPro; IPR000408; Reg\_chr\_condens.  
 DR InterPro; IPR000051; SAM\_bind.  
 DR InterPro; IPR001031; Thioesterase.  
 DR Pfam; PF01648; ACPS; 1.  
 DR Pfam; PF00975; Thioesterase; 1.  
 DR PROSITE; PS00626; RCC1\_2; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 832 AA; 89947 MW; 8F679F2A62685FCA CRC64;

Query Match 64.1%; Score 41; DB 16; Length 832;  
 Best Local Similarity 63.6%; Pred. No. 91;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAA 12  
 :|:|:|:|:|  
 Db 751 RFLAAWSLKEA 761

# RESULT 5

## Q98CR0

ID Q98CR0 PRELIMINARY; PRT; 275 AA.  
 AC Q98CR0;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Enoyl-CoA hydratase (EC 4.2.1.17).  
 GN MLR5044.  
 OS *Rhizobium loti* (*Mesorhizobium loti*).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Phyllobacteriaceae; *Mesorhizobium*.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003005; BAB51561.1; -.  
 DR GO; GO:0004300; F:enoyl-CoA hydratase activity; IEA.  
 DR GO; GO:0016508; F:long-chain-enoyl-CoA hydratase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001753; EnCoA\_hydratase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.  
 SQ SEQUENCE 275 AA; 30148 MW; 9C52AE72609AC7D3 CRC64;

Query Match 62.5%; Score 40; DB 16; Length 275;  
 Best Local Similarity 72.7%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 75 FCAGWDLKAAA 85

# RESULT 6

## Q9ETG0

ID Q9ETG0 PRELIMINARY; PRT; 422 AA.  
 AC Q9ETG0;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN ORF49.  
 OS Corynebacterium equi (Rhodococcus equi).  
 OG Plasmid pREAT701 (p33701), and Plasmid virulence.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Nocardiaceae; Rhodococcus.  
 OX NCBI\_TaxID=43767;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC33701; PLASMID=pREAT701 (p33701);  
 RA Takai S., Sekizaki T., Kakuda T., Nakamura M., Suzuki K., Ogino N.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC33701, and 103; PLASMID=pREAT701 (p33701), and virulence;  
 RX PubMed=11083803;  
 RA Takai S., Hines S.A., Sekizaki T., Nicholson V.M., Alperin D.A.,  
 RA Osaki M., Takamatsu D., Nakamura M., Suzuki K., Ogino N., Kakuda T.,  
 RA Dan H., Prescott J.F.;  
 RT "DNA sequence and comparison of virulence plasmids from Rhodococcus  
 RT equi ATCC 33701 and 103.";  
 RL Infect. Immun. 68:6840-6847(2000).  
 DR EMBL; AP001204; BAB16658.1; -.  
 DR EMBL; AF116907; AAG21752.1; -.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 422 AA; 42203 MW; D908E81B9151558B CRC64;

Query Match 62.5%; Score 40; DB 2; Length 422;  
Best Local Similarity 66.7%; Pred. No. 69;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
||||:| |||  
Db 329 KFVASWLAAAAA 340

RESULT 7

Q7TZ25

ID Q7TZ25 PRELIMINARY; PRT; 147 AA.  
AC Q7TZ25;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Possible transmembrane protein.  
GN MB2107C.  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AF2122/97;  
RX MEDLINE=22709107; PubMed=12788972;  
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
DR EMBL; BX248341; CAD96960.1; -.  
KW Complete proteome.  
SQ SEQUENCE 147 AA; 14295 MW; 4136FB14521C33A9 CRC64;

Query Match 60.9%; Score 39; DB 16; Length 147;  
Best Local Similarity 63.6%; Pred. No. 36;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
||| || :||:  
Db 10 FVAIWTARAAS 20

RESULT 8

Q7YTL3

ID Q7YTL3 PRELIMINARY; PRT; 285 AA.  
AC Q7YTL3;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein M199.7.  
GN M199.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Basham V.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; Z81104; CAE17892.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 285 AA; 31985 MW; 32AA9B1A70A07B57 CRC64;

Query Match 60.9%; Score 39; DB 5; Length 285;  
 Best Local Similarity 77.8%; Pred. No. 70;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFVAAWTLK 10  
 || ||||:!  
 Db 75 KFKAATVK 83

# RESULT 9

## Q8STE0

ID Q8STE0 PRELIMINARY; PRT; 331 AA.  
 AC Q8STE0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein ECU01\_0080 (Hypothetical protein ECU01\_1530)  
 DE (Hypothetical protein ECU02\_1560) (Hypothetical protein ECU04\_0090)  
 DE (Hypothetical protein ECU08\_0010) (Hypothetical protein  
 DE ECU08\_2090).  
 GN ECU01\_0080 OR ECU01\_1530 OR ECU02\_1560 OR ECU04\_0090 OR ECU08\_0010 OR  
 GN ECU08\_2090.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RA Genoscope;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite



RT Encephalitozoon cuniculi.";  
 RL Nature 414:450-453(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RA Genoscope;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL391737; CAD24880.1; -.  
 DR EMBL; AL391737; CAD25024.1; -.  
 DR EMBL; AL590442; CAD25185.1; -.  
 DR EMBL; AL590444; CAD25196.1; -.  
 DR EMBL; AL590448; CAD26306.1; -.  
 DR EMBL; AL590448; CAD26511.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 331 AA; 37888 MW; 5F54620501D4DF54 CRC64;

Query Match 60.9%; Score 39; DB 5; Length 331;  
 Best Local Similarity 66.7%; Pred. No. 81;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
 : ||||| |||  
 Db 3 RIYAAWTLVAAA 14

# RESULT 10

## Q8STA9

ID Q8STA9 PRELIMINARY; PRT; 612 AA.  
 AC Q8STA9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein ECU11\_0050 (Hypothetical protein  
 DE ECU05\_1680).  
 GN ECU11\_0050 OR ECU05\_1680.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RA Genoscope;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 RT Encephalitozoon cuniculi.";  
 RL Nature 414:450-453(2001).  
 DR EMBL; AL590450; CAD25915.1; -.  
 DR EMBL; AL590445; CAD26688.1; -.  
 KW Hypothetical protein.

SQ SEQUENCE 612 AA; 70929 MW; 924DE330BA373593 CRC64;

Query Match 60.9%; Score 39; DB 5; Length 612;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
: ||||| |||  
Db 3 RIYAAWTLVAAA 14

RESULT 11

Q8DFQ8

ID Q8DFQ8 PRELIMINARY; PRT; 758 AA.  
AC Q8DFQ8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Fe2+ transport system protein B.  
GN VV10148.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
RA Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AE016797; AAO08686.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0015093; F:ferrous iron transporter activity; IEA.  
DR GO; GO:0015684; P:ferrous iron transport; IEA.  
DR InterPro; IPR003373; FeoB.  
DR InterPro; IPR006073; GTP1\_OBG.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF02421; FeoB; 1.  
DR PRINTS; PR00326; GTP1OBG.  
DR TIGRFAMs; TIGR00437; feoB; 1.  
DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
KW Complete proteome.  
SQ SEQUENCE 758 AA; 82450 MW; 3ABE99077B86A57D CRC64;

Query Match 60.9%; Score 39; DB 16; Length 758;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12  
|:|:| ||: |  
Db 693 ARFIAVWTMALA 704

RESULT 12

Q93DU6

ID Q93DU6 PRELIMINARY; PRT; 96 AA.

AC Q93DU6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS unidentified bacterium.  
 OC Bacteria; environmental samples.  
 OX NCBI\_TaxID=2338;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21536188; PubMed=11679351;  
 RA Stokes H.W., Holmes A.J., Nield B.S., Holley M.P., Nevalainen K.M.H.,  
 RA Mabbutt B.C., Gillings M.R.;  
 RT "Gene Cassette PCR: Sequence-Independent Recovery of Entire Genes from  
 RT Environmental DNA.";  
 RL Appl. Environ. Microbiol. 67:5240-5246(2001).  
 DR EMBL; AF378540; AAK56385.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 96 AA; 10602 MW; 80D289940D404593 CRC64;

Query Match 59.4%; Score 38; DB 2; Length 96;  
 Best Local Similarity 50.0%; Pred. No. 35;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKAA 12  
 |:|:| | | :|  
 Db 52 ARFIATWRLSSA 63

# RESULT 13

Q8YBT0

ID Q8YBT0 PRELIMINARY; PRT; 352 AA.  
 AC Q8YBT0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative transmembrane protein.  
 GN BMEII0806.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / Biotype 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyrpides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL; AE009714; AAL54048.1; -.  
 DR PIR; AE3610; AE3610.  
 DR InterPro; IPR002549; UPF0118.

DR Pfam; PF01594; UPF0118; 1.

KW Complete proteome.

SQ SEQUENCE 352 AA; 39062 MW; 6D7CDA913601125F CRC64;

Query Match 59.4%; Score 38; DB 16; Length 352;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11

| |:| |:| |

Db 331 AMFIAAWSLLA 341

#### RESULT 14

Q8FWJ3

ID Q8FWJ3 PRELIMINARY; PRT; 352 AA.

AC Q8FWJ3;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Membrane protein, putative.

GN BRA0460.

OS Brucella suis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI\_TaxID=29461;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;

RX MEDLINE=22247741; PubMed=12271122;

RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,

RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,

RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,

RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between

RT animal and plant pathogens and symbionts.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).

DR EMBL; AE014544; AAN33654.1; -.

DR TIGR; BRA0460; -.

DR InterPro; IPR002549; UPF0118.

DR Pfam; PF01594; UPF0118; 1.

KW Complete proteome.

SQ SEQUENCE 352 AA; 39106 MW; AE6E4B8E3613829D CRC64;

Query Match 59.4%; Score 38; DB 16; Length 352;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKFVAAWTLKA 11

| |:| |:| |

Db 331 AMFIAAWSLLA 341

#### RESULT 15

039493

ID O39493 PRELIMINARY; PRT; 438 AA.  
AC O39493;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Glycoprotein M.  
GN UL10.  
OS Bovine herpesvirus type 1 (strain Jura).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=31518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Jura;  
RX MEDLINE=95313343; PubMed=7793062;  
RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,  
RA Letchworth G.J., Schwyzer M.;  
RT "Nucleotide sequence analysis of a 30-kb region of the bovine  
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with  
RT the UL21 to UL4 genes of herpes simplex virus.";  
RL Virology 210:100-108(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Jura;  
RA Schwyzer M., Paces V., Letchworth G.J., Misra V., Buhk H.J.,  
RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vlcek C.;  
RT "Complete DNA sequence of bovine herpesvirus 1.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ004801; CAA06124.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR000785; Herpes\_glycop.  
DR Pfam; PF01528; Herpes\_glycop; 1.  
DR PRINTS; PR00333; HSVINTEGRIMP.  
SQ SEQUENCE 438 AA; 45517 MW; 4E7C7FA64FAAEFC7 CRC64;

Query Match 59.4%; Score 38; DB 12; Length 438;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VAAWTLKAA 12  
:||||:|  
Db 141 LAAWTLQAA 149

Search completed: August 13, 2004, 09:15:42  
Job time : 4.98413 secs

OM protein - protein search, using sw model

Run on: August 13, 2004, 09:07:28 ; Search time 0.476191 Seconds  
 (without alignments)  
 1421.515 Million cell updates/sec

Title: US-09-785-215-19  
 Perfect score: 64  
 Sequence: 1 AKFVAAWTLKAAA 13

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	40	62.5	260	1	CAID_PROSL	Q8gb17 proteus sp.
2	39	60.9	146	1	YK81_MYCTU	Q10689 mycobacteri
3	39	60.9	260	1	CAID_ECO57	Q8xa35 escherichia
4	39	60.9	260	1	CAID_ECOL6	Q8fla6 escherichia
5	39	60.9	260	1	CAID_ECOLI	P31551 escherichia
6	39	60.9	260	1	CAID_SALTI	Q8z915 salmonella
7	39	60.9	260	1	CAID_SALTY	Q8zrx5 salmonella
8	39	60.9	260	1	CAID_SHIFL	P59395 shigella fl
9	38	59.4	155	1	LACB_MACGI	P11944 macropus gi
10	38	59.4	174	1	LACB_MACEU	Q29614 macropus eu
11	38	59.4	204	1	UBCX_PICPA	P49428 pichia past
12	38	59.4	411	1	VGLM_HSVBC	P52370 bovine herp
13	37	57.8	410	1	APGM_PYRAB	Q9v2m6 pyrococcus
14	37	57.8	411	1	APGM_PYRFU	P58814 pyrococcus
15	36	56.2	277	1	PSB2_PSESM	Q87u31 pseudomonas
16	36	56.2	362	1	YAHA_ECOLI	P21514 escherichia
17	36	56.2	399	1	DXR_BORBR	Q7wj88 bordetella

18	36	56.2	399	1	DXR_BORPA	Q7wa54	bordetella
19	36	56.2	399	1	DXR_BORPE	Q7vyc4	bordetella
20	36	56.2	776	1	TFR1_CHICK	Q90997	gallus gall
21	35	54.7	128	1	U384_HSVMG	Q05102	marek's dis
22	35	54.7	313	1	OZA4_HUMAN	Q8ngh8	homo sapien
23	35	54.7	314	1	OZA1_HUMAN	Q8ngh5	homo sapien
24	35	54.7	550	1	PTR2_HUMAN	P49190	homo sapien
25	35	54.7	639	1	V70K_PLRVW	P11622	potato leaf
26	35	54.7	730	1	EF2_METAC	Q8trc3	methanosarc
27	35	54.7	730	1	EF2_METMA	Q8pur7	methanosarc
28	35	54.7	730	1	EF2_METTE	O93640	methanosarc
29	35	54.7	750	1	FOH1_HUMAN	Q04609	h glutamate
30	35	54.7	976	1	HMDH_GIBFU	Q12577	gibberella
31	35	54.7	1036	1	Y946_ARCFU	O29316	archaeoglob
32	35	54.7	1068	1	HMDH_ASPT	Q9y7d2	aspergillus
33	34	53.1	83	1	HPIS_THETI	P80176	thermochrom
34	34	53.1	85	1	HPIS_THIRO	P00261	thiocapsa r
35	34	53.1	122	1	HPIS_CHRVI	P00260	chromatium
36	34	53.1	379	1	SGAA_METEX	P55819	methylobact
37	34	53.1	503	1	C726_ARATH	O65787	arabidopsis
38	34	53.1	515	1	TTC8_MOUSE	Q8vd72	mus musculu
39	34	53.1	531	1	TTC8_HUMAN	Q8tam2	homo sapien
40	34	53.1	574	1	KCN2_MOUSE	P58390	mus musculu
41	34	53.1	579	1	KCN2_HUMAN	Q9h2s1	homo sapien
42	34	53.1	580	1	KCN2_RAT	P70604	rattus norv
43	34	53.1	668	1	PGTB_SALTY	P37433	salmonella
44	34	53.1	711	1	HPPA_RHIME	Q8vrz3	rhizobium m
45	34	53.1	739	1	CATA_MYCSM	Q59557	mycobacteri

# ALIGNMENTS

## RESULT 1

### CAID\_PROSL

ID CAID\_PROSL STANDARD; PRT; 260 AA.  
AC Q8GB17;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
DE hydratase).  
GN CAID.  
OS Proteus sp. (strain LE138).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Proteus.  
OX NCBI\_TaxID=217617;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Engemann C., Elssner T., Pfeifer S., Krumbholz C., Maier T.,  
RA Kleber H.-P.;  
RT "Cai locus and corresponding enzymes of Proteus sp.";  
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE OF 1-20, AND SUBUNIT.  
RX MEDLINE=21302209; PubMed=11409545;  
RA Engemann C., Elssner T., Kleber H.-P.;

RT "Biotransformation of crotonobetaine to L(-)-carnitine in Proteus  
 RT sp.";  
 RL Arch. Microbiol. 175:353-359(2001).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-  
 CC butyrobetaine); second step.  
 CC -!- SUBUNIT: Homotrimer (Probable).  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ508908; CAD48582.1; -.  
 DR HAMAP; MF\_01051; -; 1.  
 DR InterPro; IPR001753; EnCoA\_hydrase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase.  
 FT INIT\_MET 0 0  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 28053 MW; A6375AE0BD12CE43 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 5.5;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FVAAWTLKAAA 13  
 | | | | | | | |  
 Db 59 FCAGWDLKAAA 69

#### RESULT 2

YK81\_MYCTU

ID YK81\_MYCTU STANDARD; PRT; 146 AA.  
 AC Q10689;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein Rv2081c/MT2143.  
 GN RV2081C OR MT2143 OR MTCY49.20C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;



RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RX MEDLINE=22206494; PubMed=12218036;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;

RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";

RL J. Bacteriol. 184:5479-5490(2002).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; Z73966; CAA98193.1; -.

DR EMBL; AE007063; AAK46425.1; ALT\_INIT.

DR PIR; D70766; D70766.

DR TIGR; MT2143; -.

DR TubercuList; Rv2081c; -.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 1 21 POTENTIAL.

FT TRANSMEM 35 55 POTENTIAL.

FT TRANSMEM 87 107 POTENTIAL.

FT TRANSMEM 111 131 POTENTIAL.

SQ SEQUENCE 146 AA; 14238 MW; BCA294986F8C9FD2 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 146;

Best Local Similarity 63.6%; Pred. No. 4.9;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13

||| || :||:

Db 10 FVAIWTARAAS 20

RESULT 3

CAID\_ECO57

ID CAID\_ECO57 STANDARD; PRT; 260 AA.

AC Q8XA35;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE CarnitinyI-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 DE hydratase).  
 GN CAID OR Z0042 OR ECS0039.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=833334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitinyI-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitinyI-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to  
 CC gamma-butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE005180; AAG54339.1; ALT\_INIT.  
 DR EMBL; AP002550; BAB33462.1; ALT\_INIT.  
 DR HAMAP; MF\_01051; -; 1.  
 DR InterPro; IPR001753; EnCoA\_hydrase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.

FT INIT\_MET 0 0 BY SIMILARITY.  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 28006 MW; 7E1E9CC3C9290ECF CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 59 FSAGWDLKAAA 69

#### RESULT 4

##### CAID\_ECOL6

ID CAID\_ECOL6 STANDARD; PRT; 260 AA.  
 AC Q8FLA6;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 DE hydratase).  
 GN CAID OR C0045.  
 OS Escherichia coli O6.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=217992;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
 RX MEDLINE=22388234; PubMed=12471157;  
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 RT of uropathogenic Escherichia coli.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA + H(2)O = crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-  
 CC butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE016755; AAN78543.1; ALT\_INIT.  
 DR HAMAP; MF\_01051; -; 1.

DR InterPro; IPR001753; EnCoA\_hydratase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 28045 MW; F33409DD2628B3B6 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 59 FSAGWDLKAAA 69

# RESULT 5

## CAID\_ECOLI

ID CAID\_ECOLI STANDARD; PRT; 260 AA.  
 AC P31551; P75623;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 hydratase).  
 GN CAID OR B0036.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=92334977; PubMed=1630901;  
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 0-2.4 min region.";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=O44:K74;  
 RX MEDLINE=95115548; PubMed=7815937;  
 RA Eichler K., Bourgis F., Buchet A., Kleber H.-P.,  
 RA Mandrand-Berthelot M.-A.;  
 RT "Molecular characterization of the cai operon necessary for carnitine  
 RT metabolism in Escherichia coli.";  
 RL Mol. Microbiol. 13:775-786(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [4]  
 RP SEQUENCE OF 1-10, AND CHARACTERIZATION.  
 RC STRAIN=O44:K74;  
 RX MEDLINE=21435666; PubMed=11551212;  
 RA Elssner T., Engemann C., Baumgart K., Kleber H.-P.;  
 RT "Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA  
 RT hydratase and a CoA-transferase, in the hydration of crotonobetaine  
 RT to L-carnitine by Escherichia coli.";  
 RL Biochemistry 40:11140-11148(2001).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA  
 CC to crotonobetainyl-CoA.  
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to  
 CC gamma-butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D10483; BAB96605.1; ALT\_INIT.  
 DR EMBL; X73904; CAA52114.1; ALT\_INIT.  
 DR EMBL; AE000114; AAC73147.1; ALT\_INIT.  
 DR HSSP; P14604; 2DUB.  
 DR EcoGene; EG11557; caiD.  
 DR HAMAP; MF\_01051; -, 1.  
 DR InterPro; IPR001753; EnCoA\_hydrase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.  
 FT INIT\_MET 0 0  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 FT VARIANT 246 246 P -> L (IN STRAIN O44:K74).  
 FT CONFLICT 108 108 G -> A (IN REF. 1).  
 SQ SEQUENCE 260 AA; 28059 MW; 0DC5E6C3C928A5CC CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 59 FSAGWDLKAAA 69

RESULT 6  
 CAID\_SALTI  
 ID CAID\_SALTI STANDARD; PRT; 260 AA.

AC Q8Z9L5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 DE hydratase).  
 GN CAID OR STY0080 OR T0071.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ty2 / ATCC 700931;  
 RX MEDLINE=22531367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;  
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
 RT and CT18.";  
 RL J. Bacteriol. 185:2330-2337(2003).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to  
 CC gamma-butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL627265; CAD01224.1; -.  
 DR EMBL; AE016834; AAO67804.1; -.  
 DR HAMAP; MF\_01051; -; 1.  
 DR InterPro; IPR001753; EnCoA\_hydrase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.

KW Lyase; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 27945 MW; CE38ED88106A6AA5 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 59 FSAGWDLKAAA 69

# RESULT 7

## CAID\_SALTY

ID CAID\_SALTY STANDARD; PRT; 260 AA.  
 AC Q8ZRX5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE CarnitinyL-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 DE hydratase).  
 GN CAID OR STM0070.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitinyL-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitinyL-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to  
 CC gamma-butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE008696; AAL19034.1; -.  
 DR StyGene; SG?????; caid.  
 DR HAMAP; MF\_01051; -; 1.  
 DR InterPro; IPR001753; EnCoA\_hydratase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 27975 MW; 30C90296FF6A7BAF CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | |  
 Db 59 FSAGWDLKAAA 69

# RESULT 8

## CAID\_SHIFL

ID CAID\_SHIFL STANDARD; PRT; 260 AA.  
 AC P59395;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Carnitiny-CoA dehydratase (EC 4.2.1.-) (Crotonobetainyl-CoA  
 DE hydratase).  
 GN CAID OR SF0033 OR S0035.  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigella.  
 OX NCBI\_TaxID=623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 RT through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;  
 RX MEDLINE=22590274; PubMed=12704152;  
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
 RA Schwartz D.C., Blattner F.R.;  
 RT "Complete genome sequence and comparative genomics of Shigella  
 RT flexneri serotype 2a strain 2457T.";



RL Infect. Immun. 71:2775-2786(2003).  
 CC -!- FUNCTION: Catalyzes the reversible dehydration of L-carnitiny-CoA  
 CC to crotonobetainyl-CoA (By similarity).  
 CC -!- CATALYTIC ACTIVITY: L-carnitiny-CoA = H(2)O + crotonobetainyl-  
 CC CoA.  
 CC -!- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-  
 CC butyrobetaine); second step.  
 CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE015040; AAN41699.1; ALT\_INIT.  
 DR EMBL; AE016978; AAP15580.1; ALT\_INIT.  
 DR HAMAP; MF\_01051; -; 1.  
 DR InterPro; IPR001753; EnCoA\_hydrase.  
 DR Pfam; PF00378; ECH; 1.  
 DR PROSITE; PS00166; ENOYL\_COA\_HYDRATASE; 1.  
 KW Lyase; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT ACT\_SITE 110 110 BASE (BY SIMILARITY).  
 FT ACT\_SITE 130 130 ACID (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 28089 MW; 0DC5FC696839A5CC CRC64;

Query Match 60.9%; Score 39; DB 1; Length 260;  
 Best Local Similarity 72.7%; Pred. No. 8.4;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FVAAWTLKAAA 13  
 | | | | | | |  
 Db 59 FSAGWDLKAAA 69

# RESULT 9

## LACB\_MACGI

ID LACB\_MACGI STANDARD; PRT; 155 AA.  
 AC P11944;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Beta-lactoglobulin.  
 GN LGB.  
 OS Macropus giganteus (Eastern gray kangaroo).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 OX NCBI\_TaxID=9317;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=87299024; PubMed=3620116;  
 RA Godovac-Zimmermann J., Shaw D.;  
 RT "Beta-lactoglobulin identified in marsupial milk. The primary  
 RT structure, binding site and possible function of beta-lactoglobulin

RT from eastern grey kangaroo (*Macropus giganteus*).";  
 RL Biol. Chem. Hoppe-Seyler 368:879-886(1987).  
 CC -!- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT  
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF  
 CC THAT MOLECULE.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SIMILARITY: Belongs to the lipocalin family.  
 DR PIR; A29699; A29699.  
 DR HSSP; P02754; 1BSQ.  
 DR InterPro; IPR000566; Lipocln\_cytFABP.  
 DR Pfam; PF00061; lipocalin; 1.  
 DR PROSITE; PS00213; LIPOCALIN; FALSE\_NEG.  
 KW Milk; Whey; Retinol-binding; Transport; Lipocalin.  
 FT DISULFID 61 153 BY SIMILARITY.  
 FT DISULFID 103 115 BY SIMILARITY.  
 SQ SEQUENCE 155 AA; 18340 MW; 7BD7EA4A191530E5 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 155;  
 Best Local Similarity 58.3%; Pred. No. 7.8;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
 ||| :| |: ||  
 Db 14 KFGVSWYLREAA 25

#### RESULT 10

##### LACB\_MACEU

ID LACB\_MACEU STANDARD; PRT; 174 AA.  
 AC Q29614;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Beta-lactoglobulin precursor.  
 GN LGB OR BLG.  
 OS *Macropus eugenii* (Tammur wallaby).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; *Macropus*.  
 OX NCBI\_TaxID=9315;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91197370; PubMed=1707629;  
 RA Collet C., Joseph R., Nicholas K.R.;  
 RT "A marsupial beta-lactoglobulin gene: characterization and prolactin-  
 RT dependent expression.";  
 RL J. Mol. Endocrinol. 6:9-16(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95314595; PubMed=7794241;  
 RA Collet C., Joseph R.;  
 RT "Exon organization and sequence of the genes encoding alpha-  
 RT lactalbumin and beta-lactoglobulin from the tammur wallaby  
 RT (*Macropodidae*, *Marsupialia*).";  
 RL Biochem. Genet. 33:61-72(1995).  
 RN [3]  
 RP SEQUENCE OF 47-174 FROM N.A.  
 RC TISSUE=Mammary gland;

RX MEDLINE=95085593; PubMed=7993373;  
 RA Collet C., Joseph R.;  
 RT "The identification of nuclear and mitochondrial genes by sequencing  
 RT randomly chosen clones from a marsupial mammary gland cDNA library."  
 RL Biochem. Genet. 32:181-190(1994).  
 CC -!- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT  
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF  
 CC THAT MOLECULE.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SIMILARITY: Belongs to the lipocalin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L14959; AAA31597.1; -.  
 DR EMBL; L14954; AAA31597.1; JOINED.  
 DR EMBL; L14955; AAA31597.1; JOINED.  
 DR EMBL; L14956; AAA31597.1; JOINED.  
 DR EMBL; L14957; AAA31597.1; JOINED.  
 DR EMBL; L14958; AAA31597.1; JOINED.  
 DR EMBL; X15212; CAA33282.1; -.  
 DR HSSP; P02754; 1BSY.  
 DR InterPro; IPR000566; Lipocln\_cytFABP.  
 DR Pfam; PF00061; lipocalin; 1.  
 DR PROSITE; PS00213; LIPOCALIN; FALSE\_NEG.  
 KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 174 BETA-LACTOGLOBULIN.  
 FT DISULFID 79 172 BY SIMILARITY.  
 FT DISULFID 122 134 BY SIMILARITY.  
 FT CONFLICT 76 76 A -> T (IN REF. 3).  
 SQ SEQUENCE 174 AA; 20219 MW; 6C5901051F99C991 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 174;  
 Best Local Similarity 58.3%; Pred. No. 8.7;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
 ||| :| |: ||  
 Db 32 KFGGSWYLREAA 43

# RESULT 11

## UBCX\_PICPA

ID UBCX\_PICPA STANDARD; PRT; 204 AA.  
 AC P49428;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2-24 kDa (EC 6.3.2.19) (Ubiquitin-  
 DE protein ligase) (Ubiquitin carrier protein) (Peroxin-4).  
 GN PEX4 OR PAS4.

OS *Pichia pastoris* (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; *Pichia*.  
 OX NCBI\_TaxID=4922;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94342381; PubMed=8063827;  
 RA Crane D.I., Kalish J.E., Gould S.J.;  
 RT "The *Pichia pastoris* PAS4 gene encodes a ubiquitin-conjugating enzyme  
 RT required for peroxisome assembly."  
 RL J. Biol. Chem. 269:21835-21844(1994).  
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO  
 CC OTHER PROTEINS. ESSENTIAL FOR PEROXISOME BIOGENESIS.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- PATHWAY: Ubiquitin conjugation; second step.  
 CC -!- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC STRONGEST, TO YEAST UBC10.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U12511; AAA53634.1; -.  
 DR PIR; A53848; A53848.  
 DR HSSP; Q95044; 2E2C.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE; PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ubl conjugation pathway; Ligase; Multigene family; Peroxisome.  
 FT BINDING 133 133 UBIQUITIN (BY SIMILARITY).  
 FT MUTAGEN 133 133 C->S,A: LOSS OF ACTIVITY.  
 SQ SEQUENCE 204 AA; 23565 MW; 46680ABD73121F6C CRC64;

Query Match 59.4%; Score 38; DB 1; Length 204;  
 Best Local Similarity 58.3%; Pred. No. 10;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKEVAAWTLKAA 12  
 ||: |||| :|  
 Db 139 AKWTPAWTLSSA 150

RESULT 12  
 VGLM\_HSVBC  
 ID VGLM\_HSVBC STANDARD; PRT; 411 AA.  
 AC P52370;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Glycoprotein M.  
 GN GM OR UL10.  
 OS Bovine herpesvirus type 1 (strain Cooper).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=10323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95313343; PubMed=7793062;  
 RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,  
 RA Letchworth G.J., Schwyzer M.;  
 RT "Nucleotide sequence analysis of a 30-kb region of the bovine  
 RT herpesvirus 1 genome which exhibits a colinear gene arrangement with  
 RT the UL21 to UL4 genes of herpes simplex virus."  
 RL Virology 210:100-108(1995).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein M family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z48053; CAA88123.1; -.  
 DR PIR; S61245; S61245.  
 DR InterPro; IPR000785; Herpes\_glycop.  
 DR Pfam; PF01528; Herpes\_glycop; 1.  
 DR PRINTS; PR00333; HSVINTEGRIMP.  
 KW Transmembrane; Glycoprotein.  
 FT TRANSMEM 14 34 POTENTIAL.  
 FT TRANSMEM 89 109 POTENTIAL.  
 FT TRANSMEM 133 153 POTENTIAL.  
 FT TRANSMEM 159 179 POTENTIAL.  
 FT TRANSMEM 213 233 POTENTIAL.  
 FT TRANSMEM 249 269 POTENTIAL.  
 FT TRANSMEM 277 297 POTENTIAL.  
 FT TRANSMEM 318 338 POTENTIAL.  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 411 AA; 43029 MW; 20F156DA9F40158C CRC64;

Query Match 59.4%; Score 38; DB 1; Length 411;  
 Best Local Similarity 77.8%; Pred. No. 19;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VAAWTLKAA 12  
 :||||:|  
 Db 141 LAAWTLQAA 149

RESULT 13  
 APGM\_PYRAB

ID APMG\_PYRAB STANDARD; PRT; 410 AA.  
AC Q9V2M6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase  
DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGAM) (aPGAM).  
GN APMG OR PYRAB00490 OR PAB2318.  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GE5 / Orsay;  
RX MEDLINE=22511545; PubMed=12622808;  
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,  
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;  
RT "An integrated analysis of the genome of the hyperthermophilic  
RT archaeon Pyrococcus abyssi.";  
RL Mol. Microbiol. 47:1495-1512(2003).  
CC -!- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and  
CC 3-phosphoglycerate (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.  
CC -!- SIMILARITY: Belongs to the BPG-independent phosphoglycerate mutase  
CC family. A-PGAM subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ248283; CAB48972.1; -.  
DR PIR; E75190; E75190.  
DR HAMAP; MF\_01402; -, 1.  
DR InterPro; IPR004456; BcpB.  
DR InterPro; IPR006124; Metalloenzyme.  
DR Pfam; PF01676; Metalloenzyme; 1.  
DR ProDom; PD004704; BcpB; 1.  
DR TIGRFAMs; TIGR00306; bcpB; 1.  
KW Isomerase; Complete proteome.  
SQ SEQUENCE 410 AA; 45201 MW; 91C8A89637608826 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 410;  
Best Local Similarity 58.3%; Pred. No. 29;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 KFVAAWTLKAAA 13  
|| | :|||  
Db 235 KFTEQWKVKAAA 246

APMG\_PYRFU  
 ID APMG\_PYRFU STANDARD; PRT; 411 AA.  
 AC P58814;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase  
 DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGAM) (aPGAM).  
 GN APMG OR PF1959.  
 OS *Pyrococcus furiosus*.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC *Pyrococcus*.  
 OX NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the *Pyrococcus furiosus* genome.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP CHARACTERIZATION.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RX MEDLINE=22071963; PubMed=12076796;  
 RA van der Oost J., Huynen M.A., Verhees C.H.;  
 RT "Molecular characterization of phosphoglycerate mutase in archaea.";  
 RL FEMS Microbiol. Lett. 212:111-120(2002).  
 CC -!- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and  
 CC 3-phosphoglycerate.  
 CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.  
 CC -!- COFACTOR: Magnesium (Probable).  
 CC -!- ENZYME REGULATION: Inhibited to approximately 20% by EDTA.  
 CC -!- SUBUNIT: Homotetramer (Probable).  
 CC -!- MISCELLANEOUS: Optimal pH is 8.0.  
 CC -!- SIMILARITY: Belongs to the BPG-independent phosphoglycerate mutase  
 CC family. A-PGAM subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE010289; AAL82083.1; -.  
 DR HAMAP; MF\_01402; -, 1.  
 DR InterPro; IPR004456; BcpB.  
 DR InterPro; IPR006124; Metalloenzyme.  
 DR Pfam; PF01676; Metalloenzyme; 1.  
 DR ProDom; PD004704; BcpB; 1.  
 DR TIGRFAMs; TIGR00306; bcpB; 1.  
 KW Isomerase; Magnesium; Complete proteome.  
 SQ SEQUENCE 411 AA; 45314 MW; 870587E630C2B104 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 411;  
 Best Local Similarity 58.3%; Pred. No. 29;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFVAAWTLKAAA 13  
|| | :|||  
Db 236 KFTEQWKVKAAA 247

RESULT 15

PSB2\_PSESM

ID PSB2\_PSESM STANDARD; PRT; 277 AA.

AC Q87U31;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Phosphate import ATP-binding protein pstB 2 (EC 3.6.3.27) (Phosphate-  
DE transporting ATPase 2) (ABC phosphate transporter 2).

GN PSTB2 OR PSTB-2 OR PSPTO5484.

OS *Pseudomonas syringae* (pv. tomato).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; *Pseudomonas*.

OX NCBI\_TaxID=323;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DC3000;

RX MEDLINE=22834015; PubMed=12928499;

RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,

RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,

RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,

RA Nelson W.C., Davidsen T., Zafar N., Zhou L., Liu J., Yuan Q.,

RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,

RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,

RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,

RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,

RA White O., Fraser C.M., Collmer A.;

RT "The complete genome sequence of the Arabidopsis and tomato pathogen

RT *Pseudomonas syringae* pv. tomato DC3000.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).

CC -!- FUNCTION: Part of the ABC transporter complex pstSACB

CC (TC 3.A.1.7.1) involved in phosphate import. Responsible for  
CC energy coupling to the transport system (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + phosphate(Out) = ADP + phosphate  
CC + phosphate(In).

CC -!- SUBUNIT: The complex is composed of two ATP-binding proteins  
CC (pstB), two transmembrane proteins (pstC and pstA) and a solute-  
CC binding protein (pstS) (Probable).

CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).

CC -!- SIMILARITY: Belongs to the ABC transporter family. PstB subfamily.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE016876; AA058903.1; -.

DR TIGR; PSPTO5484; -.



DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KW Hydrolase; Transport; Phosphate transport; Membrane; Inner membrane;  
KW ATP-binding; Complete proteome.  
FT NP\_BIND 63 70 ATP (By similarity).  
SQ SEQUENCE 277 AA; 31133 MW; E0FEE61769458723 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 277;  
Best Local Similarity 87.5%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AWTLKAAA 13  
|| |||||  
Db 151 AWALKAAA 158

Search completed: August 13, 2004, 09:13:21  
Job time : 2.47619 secs